



TUGAS AKHIR – SS141501

***BOOTSTRAP AGGREGATING MULTIVARIATE
ADAPTIVE REGRESSION SPLINES
(BAGGING MARS) UNTUK KLASIFIKASI
PASIEN HIV/AIDS DI KABUPATEN PASURUAN***

**LYYIN NAHRIYAH
NRP 1315 105 024**

**Dosen Pembimbing
Dr. Bambang Widjanarko Otok, M.Si**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2017**



TUGAS AKHIR – SS141501

***BOOTSTRAP AGGREGATING MULTIVARIATE
ADAPTIVE REGRESSION SPLINES
(BAGGING MARS) UNTUK KLASIFIKASI
PASIEN HIV/AIDS DI KABUPATEN PASURUAN***

**LYYIN NAHRIYAH
NRP 1315 105 024**

**Dosen Pembimbing
Dr. Bambang Widjanarko Otok, M.Si**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2017**



FINAL PROJECT – SS141501

**BOOTSTRAP AGGREGATING MULTIVARIATE
ADAPTIVE REGRESSION SPLINES
(BAGGING MARS) FOR CLASSIFICATION
HIV/AIDS PATIENTS IN PASURUAN**

**LYYIN NAHRIYAH
NRP 1315 105 024**

**Supervisor
Dr. Bambang Widjanarko Otok, M.Si**

**UNDERGRADUATE PROGRAM
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS AND NATURAL SCIENCES
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2017**

LEMBAR PENGESAHAN

BOOTSTRAP AGGREGATING MULTIVARIATE ADAPTIVE REGRESSION SPLINES (BAGGING MARS) UNTUK KLASIFIKASI PASIEH HIV/AIDS DI KABUPATEN PASURUAN

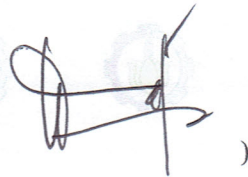
TUGAS AKHIR

Diajukan Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Sains
pada

Program Studi Sarjana Departemen Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember

Oleh :

Lyyin Nahriyah
NRP. 1315 105 024



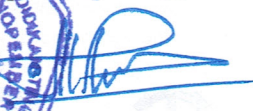
Disetujui oleh Pembimbing:

Dr. Bambang Widjanarko Otok, M.Si
NIP. 19681124 199412 1 001

()



Mengetahui,
Kepala Departemen


Dr. Suhartono
NIP. 19710929 199512 1 001

SURABAYA, JULI 2017

**BOOTSTRAP AGGREGATING MULTIVARIATE
ADAPTIVE REGRESSION SPLINES
(BAGGING MARS) UNTUK KLASIFIKASI PASIEN
HIV/AIDS DI KABUPATEN PASURUAN**

Nama Mahasiswa : Lyyin Nahriyah
NRP : 1315 105 024
Departemen : Statistika
Dosen Pembimbing : Dr. Bambang Widjanarko Otok, M.Si

Abstrak

HIV merupakan virus yang menyerang sel darah putih, sedangkan penyakit AIDS adalah kumpulan gejala penyakit akibat menurunnya sistem kekebalan tubuh yang disebabkan oleh virus HIV. Diagnosis HIV secara dini dan penanganan yang efektif, pengidap HIV tidak akan berubah menjadi AIDS. Penyakit AIDS sangat berbahaya karena dapat menyebabkan kematian sebagai akibat langsung dari salah satu atau lebih adanya Infeksi Oportunistik (IO) dan diperkirakan angka orang dengan HIV dan AIDS di Indonesia akan semakin meningkat bila tidak ada pencegahan. Pemodelan kejadian dan ketepatan klasifikasi kejadian IO dari faktor-faktor yang mempengaruhi diperlukan, agar dapat mempersiapkan segala kemungkinan yang akan terjadi dan pencegahannya, menggunakan metode Bootstrap Aggregating Multivariate Adaptive Regression Splines (Bagging MARS). Data yang digunakan merupakan data sekunder yang berasal dari rekam medis pasien HIV/AIDS di Puskesmas yang ada di Kabupaten Pasuruan. Model MARS yang telah terbentuk didapatkan probabilitas tidak terjadinya kasus Infeksi Oportunistik pada pasien sebesar 0,717 dan terjadinya IO adalah sebesar 0,283. Perbandingan antara metode MARS dan Bagging MARS didapatkan kesimpulan bahwa pada klasifikasi kejadian IO dengan data training kedua metode sama-sama menghasilkan akurasi yang terbaik, dan untuk klasifikasi pada data testing lebih tepat jika menggunakan metode Bagging MARS dengan replikasi sejumlah 50 kali.

Kata kunci: AIDS, Bagging, HIV, Infeksi Oportunistik, Klasifikasi, MARS

(Halaman ini sengaja dikosongkan)

BOOTSTRAP AGGREGATING MULTIVARIATE ADAPTIVE REGRESSION SPLINES (BAGGING MARS) FOR CLASSIFICATION HIV/AIDS PATIENTS IN PASURUAN

Student Name : Lyyin Nahriyah
NRP : 1315 105 024
Department : Statistics
Supervisor : Dr. Bambang Widjanarko Otok, M.Si

Abstract

HIV is a virus which attacks white blood cells, while AIDS is a set of disease symptoms occurring from weakened immune system caused by HIV. Early diagnosis and effective handling of the virus will help HIV patients to be less susceptible to suffering AIDS. AIDS is very lethal because it could result in fatality as a repercussion from one or more Opportunistic Infection (OI) and the number of patients with HIV/AIDS in Indonesia are predicted to rise if there are no precautions. Based on that data, a model is needed to estimate the chance of events occurring and accuracy of classification of OI events from influencing factors, and to prepare and anticipate for any possibilities that could happen by using Bootstrap Aggregating Multivariate Adaptive Regression Splines (Bagging MARS). The data used in this research is a secondary data taken from medical records of HIV/AIDS patients in Community Health Clinic of Pasuruan. Based on MARS model, it is obtained a 0,717 probability of Opportunistic Infection not occurring on patients and 0,283 chance of OI occurring, meanwhile from the comparison of both methods it is concluded that the classification of OI using training data results in best accuracy, while for testing data it is more accurate to use Bagging MARS method with 50 times replication.

Keywords : AIDS, Bagging, Classification, HIV, MARS Opportunistic Infection

(Halaman ini sengaja dikosongkan)

KATA PENGANTAR

Alhamdulillahirobbil'Alamin, puji syukur yang tiada tara kehadiran Allah SWT atas limpahan rahmat, nikmat, dan hidayah-Nya, sehingga penulis dapat menyelesaikan laporan Tugas Akhir yang berjudul ***“Bootstrap Aggregating Multivariate Adaptive Regression Splines (Bagging MARS) untuk Klasifikasi Pasien HIV/AIDS di Kabupaten Pasuruan”***. Sholawat serta salam semoga tetap terlimpahkan kepada suri tauladan terbaik bagi seluruh ummat, Rasulullah SAW dan keluarganya.

Terselesaikannya Tugas Akhir ini, tentu bukan hanya karena usaha dari penulis, tapi ada banyak pihak yang sangat berjasa dalam proses pengerjaan Tugas Akhir ini. Penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Bapak Dr. Bambang Widjanarko Otok, M.Si selaku dosen pembimbing yang dengan sabar mengajari, memberikan ilmu, saran, dan waktu yang beliau luangkan untuk membimbing penulis dalam menyelesaikan Tugas Akhir ini.
2. Puskesmas dan Dinas Kesehatan Kabupaten Pasuruan yang telah berkenan memberikan ijin pengambilan data sehingga dapat digunakan pada Tugas Akhir ini.
3. Bapak Dr. Purhadi, M.Sc dan Ibu Diaz Fitra Aksioma, S.Si, M.Si selaku dosen penguji atas ilmu, saran, masukan dan koreksinya pada Tugas Akhir ini.
4. Bapak Dr. Suhartono selaku Kepala Departemen Statistika FMIPA ITS yang telah memberikan fasilitas dan sarana dalam penyusunan Tugas Akhir ini..
5. Bapak Dr. Sutikno, M.Si selaku Ketua Program Studi S1 Departemen Statistika FMIPA ITS yang telah mengawal proses berjalannya Tugas Akhir dengan bimbingan dan fasilitas yang diberikan.
6. Seluruh Bapak-Ibu Dosen Statistika atas segala ilmu yang diberikan dan seluruh staf serta karyawan Jurusan Statistika.
7. Abah dan Ibu terkasih, tercinta, tersayang, serta mbak Cicik dan seluruh keluarga, terimakasih yang tak henti-hentinya

memberi doa dan dukungan baik moril serta materil kepada penulis sehingga mampu menyelesaikan laporan Tugas Akhir ini. Semoga senantiasa diberkahi oleh-Nya.

8. Seluruh teman-teman mahasiswa Lintas Jalur S1 Statistika ITS angkatan 2015 yang telah membantu hingga akhir perkuliahan dan Mahasiswa S1 Bimbingan Pak Bambang yang selalu memberikan bantuan dan semangat hingga terselesaikannya Tugas Akhir.
9. Dan seluruh pihak yang tidak dapat disebutkan satu per satu.

Penulis menyadari bahwa laporan ini masih belum sempurna, oleh karena itu penulis memohon maaf serta mengharapkan saran dan kritik dari pembaca agar kedepannya dapat menjadi lebih baik. Semoga laporan Tugas Akhir ini dapat memberikan manfaat bagi semua pihak yang membutuhkan.

Surabaya, Juli 2017

Penulis

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
LEMBAR PENGESAHAN	iii
ABSTRAK	v
ABSTRACT	vii
KATA PENGANTAR	ix
DAFTAR ISI	xi
DAFTAR GAMBAR	xiii
DAFTAR TABEL	xv
DAFTAR LAMPIRAN	xix
BAB I PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian	4
1.5 Batasan Masalah	4
BAB II TINJAUAN PUSTAKA	
2.1 <i>Multivariate Adaptive Regression Splines</i> (MARS) ...	5
2.2 Klasifikasi MARS Respon Biner	10
2.3 Ketepatan dan Pengujian Kestabilan Klasifikasi	11
2.4 <i>Bootstrap Aggregating</i> (Bagging)	13
2.5 Pengertian HIV/AIDS	14
2.6 Faktor yang Mempengaruhi Kejadian Infeksi Opportunistik Penderita HIV/AIDS	15
BAB III METODOLOGI PENELITIAN	
3.1 Sumber Data	19
3.2 Variabel Penelitian	19
3.3 Metode Pengolahan Data	21
3.4 Diagram Alir Penelitian	23
BAB IV ANALISIS DAN PEMBAHASAN	
4.1 Karakteristik Pasien Penderita Penyakit HIV/AIDS ..	25
4.2 Klasifikasi dan Pemodelan Kejadian IO pada Pasien HIV/AIDS Menggunakan Metode MARS	33

4.3 Klasifikasi Kejadian Infeksi Oportunistik pada Pasien HIV/AIDS dengan Metode Bagging MARS	43
4.4 Perbandingan Ketepatan Klasifikasi Kejadian IO Menggunakan Metode MARS dan Bagging MARS ..	45
BAB V KESIMPULAN DAN SARAN	
5.1 Kesimpulan	47
5.2 Saran	48
DAFTAR PUSTAKA	49
LAMPIRAN	51
BIODATA PENULIS	61

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Kerangka Konsep Teori Lawrence Perilaku Penyakit HIV/ AIDS	17
Gambar 3.2 Diagram Alir Pemodelan MARS dan Bagging MARS	23
Gambar 4.1 Persentase Kejadian Infeksi Oportunistik Pasien HIV/AIDS	26
Gambar 4.2 <i>Boxplot</i> Umur Pasien HIV/AIDS	26
Gambar 4.3 Persentase Jenis Kelamin Pasien HIV/AIDS	27
Gambar 4.4 Jumlah Pasien HIV/AIDS Berdasarkan Kriteria Pekerjaan dan Kejadian Infeksi Oportunistik	27
Gambar 4.5 Jumlah Pasien HIV/AIDS Berdasarkan Kriteria Pendidikan dan Kejadian Infeksi Oportunistik ..	28
Gambar 4.6 Jumlah Pasien HIV/AIDS Berdasarkan Kriteria Status Pernikahan	29
Gambar 4.7 <i>Boxplot</i> Lama Terjadinya Penyakit HIV/AIDS	29
Gambar 4.8 <i>Boxplot</i> Kadar CD4 pada Pasien HIV/AIDS	30
Gambar 4.9 Jumlah Pasien HIV/AIDS Berdasarkan Kriteria Perlakuan dan Kejadian IO	31
Gambar 4.10 Jumlah Pasien HIV/AIDS Berdasarkan Kriteria Pengetahuan dan Konsep Diri pada Pasien HIV/AIDS	31
Gambar 4.11 Jumlah Jumlah Pasien HIV/AIDS Berdasarkan Kriteria Dukungan Keluarga dan Kualitas Hidup Pasien HIV/AIDS	32

(Halaman ini sengaja dikosongkan)

DAFTAR TABEL

	Halaman
Tabel 2.1 Klasifikasi MARS Respon Biner.....	11
Tabel 3.1 Variabel Penelitian	19
Tabel 4.1 Kriteria Pembagian Data <i>Training</i> dan <i>Testing</i>	33
Tabel 4.2 <i>Trial</i> dan <i>Error</i> Pembentukan Model MARS	
Klasifikasi Kejadian Infeksi Oportunistik	34
Tabel 4.3 Statistik Uji Pengujian Serentak	37
Tabel 4.4 Statistik Uji Pengujian Parsial	37
Tabel 4.5 Tingkat Kepentingan Variabel Prediktor pada Model	
MARS	40
Tabel 4.6 Ketepatan Klasifikasi Data <i>Training</i> pada Model	
MARS untuk Kejadian Infeksi Oportunistik.....	41
Tabel 4.7 Statistik Uji Kestabilan Data <i>Training</i> pada Model	
MARS.....	42
Tabel 4.8 Ketepatan Klasifikasi Data <i>Testing</i> pada Model	
MARS untuk Kejadian Infeksi Oportunistik.....	43
Tabel 4.9 Statistik Uji Kestabilan Data <i>Testing</i> pada Model	
MARS.....	43
Tabel 4.10 Hasil Ketepatan Klasifikasi Data <i>Training</i> Metode	
Bagging MARS	44
Tabel 4.11 Hasil Ketepatan Klasifikasi Data <i>Testing</i> Metode	
Bagging MARS	44
Tabel 4.12 Perbandingan Hasil Ketepatan Klasifikasi Data	
<i>Training</i> Menggunakan Metode MARS dan	
Bagging MARS	45
Tabel 4.13 Perbandingan Hasil Ketepatan Klasifikasi Data	
<i>Testing</i> Menggunakan Metode MARS dan Bagging	
MARS.....	46

(Halaman ini sengaja dikosongkan)

DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Surat Pernyataan Pengambilan Data	51
Lampiran 2. Data Rekam Medis Pasien HIV/AIDS	52
Lampiran 3. Model MARS Data <i>Training</i> dengan Kombinasi BF = 24, MI = 3, dan MO = 1 menggunakan AMOS 2	53
Lampiran 4. Ketepatan Klasifikasi Data <i>Testing</i>	56
Lampiran 5. <i>Syntax</i> Program Bagging MARS pada <i>software</i> R	56
Lampiran 6. Ketepatan Klasifikasi Data <i>Training</i> dan <i>Testing</i> dengan Bagging MARS	57
Lampiran 7. <i>Trial and Error</i> Pembentukan Model MARS dengan Pembagian 80:20 pada Data	57
Lampiran 8. <i>Trial and Error</i> Pembentukan Model MARS dengan Pembagian 90:10 pada Data	59

(Halaman ini sengaja dikosongkan)

BAB I

PENDAHULUAN

1.1 Latar Belakang

Infeksi *Human Immunodeficiency Virus* (HIV) pertama kali ditemukan di Indonesia pada tahun 1987 dan kasus HIV terus mengalami peningkatan serta menduduki peringkat ketiga setelah Negara India dan Cina. Menurut data Kementerian Kesehatan Republik Indonesia sejak tahun 2005 hingga September 2015 terdapat kasus HIV sebanyak 184.929 kasus. Faktor risiko penularan HIV tertinggi adalah hubungan seks yang tidak aman pada heteroseksual yakni sebesar 46,20 persen yang didapat dari laporan layanan konseling dan tes HIV. Provinsi dengan jumlah kasus HIV tertinggi yaitu DKI Jakarta (38.464 kasus), Jawa Timur (24.104 kasus), Papua (20.147 kasus), dan Jawa Barat (17.075 kasus). Sementara kasus AIDS hingga September 2015 telah terjadi sejumlah 68.917 kasus dan berdasarkan kelompok umur persentase kasus AIDS tertinggi yakni pada usia 20-39 tahun sejumlah 61,40 persen kasus (Iradat, 2015).

Provinsi Jawa Timur sejak tahun 2003 sudah ditetapkan sebagai daerah *epidemic* terkonsentrasi dengan jumlah kasus HIV/AIDS peringkat ke-2 terbanyak dan terus menerus mengalami peningkatan sejak tahun 2010. Secara kumulatif hingga bulan Desember 2015 telah ditemukan 32.646 kasus HIV, sedangkan untuk kasus AIDS yang telah dilaporkan sebanyak 14.498 kasus, dimana 3.381 (23,3%) diantaranya sudah meninggal. Dari 38 kabupaten/kota semua sudah melaporkan adanya kasus AIDS dan berdasarkan tempat asal penderita yang paling banyak ditemukan di Kota Surabaya, Kabupaten Malang, Kabupaten Pasuruan, Gresik dan Jember (Dinas Kesehatan Provinsi Jawa Timur, 2016).

Virus HIV menyerang sistem kekebalan tubuh dan melemahkan kemampuan tubuh untuk melawan infeksi penyakit, sehingga diperlukan pengobatan agar memperlambat perkembangan penyakit dan membuat penderitanya hidup lebih lama serta dapat menjalani hidup secara normal. Diagnosis HIV secara dini dan

penanganan yang efektif, membuat pengidap HIV tidak akan berubah menjadi AIDS. Penyakit AIDS sangat berbahaya karena mempunyai *case fatality rate* 100% dalam 5 tahun atau akan meninggal kurang lebih 5 tahun setelah diagnosa, sebagai akibat langsung dari salah satu atau lebih adanya Infeksi Oportunistik. Jumlah pasien HIV/AIDS yang mengalami Infeksi Oportunistik (IO) sudah menurun secara drastis sejak adanya terapi antiretroviral (ARV). IO masih menimbulkan masalah pada pasien terbukti karena banyaknya pasien rawat inap di rumah sakit dan dapat mengancam jiwa karena dapat mempengaruhi tubuh, kulit, paru, mata, otak, dan kanker (Spiritia, 1995).

Informasi mengenai perkembangan kasus HIV/AIDS diperlukan agar didapatkan gambaran masalah sebagai salah satu pendukung dalam upaya pencegahan yang disebabkan oleh adanya Infeksi Oportunistik. Diperkirakan angka ODHA (Orang Dengan HIV dan AIDS) di Indonesia akan semakin meningkat bila tidak ada pencegahan. Pemodelan diperlukan agar didapatkan peluang kejadian dan ketepatan klasifikasi kejadian Infeksi Oportunistik dari faktor-faktor yang mempengaruhi, agar dapat mempersiapkan segala kemungkinan yang akan terjadi dan pencegahannya.

Untuk melihat pola hubungan antara variabel prediktor terhadap variabel respon dapat dilakukan melalui analisis regresi. Namun apabila tidak terdapat informasi mengenai bentuk fungsi dan tidak jelasnya pola hubungan antara variabel prediktor dengan variabel respon maka analisis dapat dilakukan menggunakan pendekatan regresi nonparametrik (Eubank, 1999). Salah satu model pendekatan regresi nonparametrik untuk klasifikasi yakni menggunakan metode *Multivariate Adaptive Regression Splines* (MARS). Metode MARS memiliki kelebihan relatif fleksibel dan inovatif untuk menyelidiki pola hubungan antara variabel tanpa asumsi terhadap bentuk fungsionalnya. MARS difokuskan untuk mengatasi permasalahan dimensi tinggi dan diskontinuitas pada data serta melibatkan banyak interaksi antar variabel (Otok, dkk., 2006). Tingkat akurasi klasifikasi model

MARS dapat ditingkatkan menggunakan *resampling*, salah satunya dengan metode *Bootstrap Aggregating* (Bagging) yang dapat memperbaiki stabilitas, meningkatkan akurasi, dan kekuatan prediktif (Breiman, 1994).

Kajian mengenai klasifikasi menggunakan metode MARS pernah dilakukan oleh Oktiva (2014) untuk mengklasifikasikan rumah tangga miskin di Kabupaten Jombang. Hasil yang didapatkan bahwa metode Bagging MARS memberikan hasil ketepatan klasifikasi lebih baik dibandingkan dengan model MARS dataset tunggal. Pada penentuan variabel terkait kasus HIV/AIDS diambil dari penelitian sebelumnya oleh Sigit, dkk (2013) mengenai pemodelan *Mixture Survival* studi kasus HIV/AIDS di klinik VCT/CST RSUP DR. Kariadi Semarang dan didapatkan hasil bahwa faktor usia, jenis kelamin, tingkat pendidikan, status kerja, status perkawinan, ARV, jumlah CD4, infeksi oportunistik, status fungsional, stadium pasien, dan kepatuhan pasien menentukan waktu kelangsungan hidup pasien HIV/AIDS. Penelitian lain dari Nidhomuddin (2014) yang menyatakan bahwa variabel usia, pekerjaan, pernah ditahan kasus NAPZA, status pernikahan, dan penggunaan jarum suntik yang steril merupakan variabel terpenting pada model MARS dengan tingkat akurasi sebesar 80,28%. Penelitian selanjutnya menggunakan metode *Bootstrap Aggregating Multivariate Adaptive Regression Splines* (Bagging MARS) untuk mengklasifikasikan kejadian HIV/AIDS berdasarkan data rekam medis pasien di Puskesmas Kabupaten Pasuruan, sehingga didapatkan model pengklasifikasian untuk kejadian Infeksi Oportunistik pada pasien HIV/AIDS melalui prediksi faktor-faktor yang mempengaruhi dan peluang kejadiannya, serta dapat diketahui tingkat akurasi klasifikasi pasien yang terkena HIV/AIDS.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka permasalahan yang akan diangkat dalam penelitian ini adalah, bagaimana karakteristik data rekam medis pasien HIV/AIDS.

Bagaimana hasil pemodelan dan peluang Infeksi Oportunistik (IO) pada ODHA dari faktor-faktor yang berpengaruh menggunakan metode MARS, serta bagaimana hasil ketepatan klasifikasi dan metode yang terbaik untuk mengklasifikasikan kejadian IO pada ODHA menggunakan metode MARS dan Bagging MARS.

1.3 Tujuan Penelitian

Tujuan dari penelitian ini diantaranya adalah didapatkan karakteristik pasien HIV/AIDS berdasarkan data rekam medis. Diketahui hasil pemodelan dan peluang Infeksi Oportunistik (IO) pada ODHA dari faktor-faktor yang berpengaruh menggunakan metode MARS. Didapatkan hasil ketepatan klasifikasi dan metode yang terbaik untuk mengklasifikasikan kejadian IO pada ODHA menggunakan metode MARS dan Bagging MARS.

1.4 Manfaat Penelitian

Manfaat yang diharapkan dapat diperoleh pada penelitian ini adalah sebagai tambahan pengetahuan mengenai salah satu metode klasifikasi nonparametrik dengan MARS dan Bagging MARS. Digunakan sebagai salah satu usaha pencegahan terjadinya kasus Infeksi Oportunistik (IO) pada ODHA melalui model prediksi yang tepat dari faktor-faktor yang berpengaruh, sehingga dapat mencegah adanya IO berat yang menyebabkan kematian.

1.5 Batasan Masalah

Batasan pada penelitian ini hanya mencakup data pengamatan pada pasien penyakit HIV/AIDS di Puskesmas Kabupaten Pasuruan. Replikasi pada metode Bagging MARS digunakan sejumlah 50, 100, 150, 200, dan 500 kali replikasi.

BAB II

TINJAUAN PUSTAKA

2.1 *Multivariate Adaptive Regression Splines (MARS)*

Metode Statistik yang paling banyak dilakukan adalah metode analisis regresi yang digunakan untuk mengetahui hubungan dan pengaruh antara variabel prediktor terhadap variabel respon, dengan variabel random *error* yang diasumsikan independen dan berdistribusi normal. Untuk melihat pola hubungan antara variabel prediktor terhadap variabel respon dapat dilakukan melalui tiga model pendekatan yakni parametrik, nonparametrik, dan semiparametrik. Namun apabila tidak terdapat informasi mengenai bentuk fungsi dan tidak jelasnya pola hubungan antara variabel prediktor dengan variabel respon maka analisis dapat dilakukan menggunakan regresi nonparametrik (Eubank, 1999).

Salah satu pendekatan regresi nonparametrik adalah *Multivariate Adaptive Regression Spline (MARS)* yang diperkenalkan oleh Friedman pada tahun 1991 dan sesuai untuk digunakan pada pola data yang tidak diketahui bentuk kurva regresinya dan atau tidak terdapat informasi masa lalu lengkap mengenai bentuk pola data (Eubank, 1999). Model MARS difokuskan untuk mengatasi permasalahan berdimensi tinggi dan diskontinuitas pada data (Hasyim, 2010). Permasalahan dimensi tinggi adalah suatu masalah dengan jumlah variabel banyak dan ukuran sampel yang besar sehingga memerlukan perhitungan yang rumit. Salah satu keunggulan menggunakan metode MARS adalah pada kemampuannya untuk memperkirakan kontribusi fungsi basis terhadap variabel respon, dengan tidak hanya dapat menangkap efek adaptif tetapi juga dapat menangkap efek interaksi antar prediktor (Friedman J. H., 1990). Metode MARS menjadi populer karena tidak menentukan tipe khusus seperti hubungan (linier, kuadratik, dan kubik) diantara variabel prediktor dan respon pada proses pembentukan model MARS tidak memerlukan asumsi (Otok, dkk., 2006).

Secara umum Model MARS menurut Friedman (1991) dapat ditulis pada persamaan (2.1) berikut:

$$y_i = \alpha_0 + \sum_{m=1}^M \alpha_m B_m(x) + \varepsilon_i \quad (2.1)$$

dengan, α_0 adalah koefisien konstanta dari basis fungsi B_0 , dan α_m merupakan koefisien dari basis fungsi ke- m , serta $B_m(\mathbf{x}) = \prod_{k=1}^{K_m} [s_{km}(x_{v(k,m)} - t_{km})]$. Sehingga jika ditulis dalam bentuk matriks adalah sebagai berikut:

$$\mathbf{y} = \mathbf{B}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}$$

dengan,

$$\mathbf{y} = (y_1, y_2, \dots, y_n)^T, \boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_M)^T, \boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)^T$$

$$\mathbf{B} = \begin{bmatrix} 1 & \prod_{K=1}^{K_1} s_{1m}(x_{1(1,m)} - t_{1m}) & \cdots & \prod_{K=1}^{K_1} s_{Mm}(x_{1(M,m)} - t_{Mm}) \\ 1 & \prod_{K=1}^{K_1} s_{1m}(x_{2(1,m)} - t_{1m}) & \cdots & \prod_{K=1}^{K_1} s_{Mm}(x_{2(M,m)} - t_{Mm}) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \prod_{K=1}^{K_1} s_{1m}(x_{n(1,m)} - t_{1m}) & \cdots & \prod_{K=1}^{K_1} s_{Mm}(x_{n(M,m)} - t_{Mm}) \end{bmatrix} \quad (2.2)$$

Modifikasi Friedman untuk mengestimasi model MARS ditulis dalam persamaan berikut:

$$\hat{f}(\mathbf{x}) = \alpha_0 + \sum_{m=1}^M \alpha_m \prod_{k=1}^{K_m} [s_{km} \cdot (x_{v(k,m)} - t_{km})]_+ \quad (2.3)$$

dengan fungsi,

$$(x_{v(k,m)} - t_{km})_+ = \begin{cases} (x_{v(k,m)} - t_{km}), & x_{v(k,m)} - t_{km} > 0 \\ 0, & x_{v(k,m)} - t_{km} \leq 0 \end{cases}$$

dimana α_0 merupakan koefisien konstanta dari basis fungsi B_0 , α_m adalah koefisien dari basis fungsi ke- m , $x_{v(k,m)}$ adalah variabel independen, t_{km} merupakan nilai knot dari variabel independen $x_{v(k,m)}$, dengan M adalah banyaknya fungsi basis, K_m adalah banyaknya interaksi pada fungsi basis ke- m . s_{km} merupakan suatu nilai yang bernilai 1 jika data berada disebelah kanan titik knot atau -1 jika data berada disebelah kiri titik knot, v merupakan

banyaknya variabel prediktor, dan k adalah banyaknya interaksi. Sehingga persamaan (2.3) dapat dijabarkan sebagai berikut.

$$\begin{aligned} \hat{f}(\mathbf{x}) = & a_0 + \sum_{m=1}^M a_m [s_{1m} \cdot (x_{v(1,m)} - t_{1m})]_+ + \sum_{m=1}^M a_m [s_{1m} \cdot (x_{v(1,m)} - t_{1m})]_+ \cdot [s_{2m} \cdot (x_{v(2,m)} - t_{2m})]_+ \\ & + \sum_{m=1}^M a_m [s_{1m} \cdot (x_{v(1,m)} - t_{1m})]_+ \cdot [s_{2m} \cdot (x_{v(2,m)} - t_{2m})]_+ \\ & \cdot [s_{3m} \cdot (x_{v(3,m)} - t_{3m})]_+ + \dots \end{aligned} \quad (2.4)$$

Secara umum persamaan (2.4) dapat dituliskan sebagai berikut:

$$\begin{aligned} \hat{f}(\mathbf{x}) = & a_0 + \sum_{i=1}^v f_i(x_i) + \sum_{\substack{i,j=1 \\ i \neq j}}^v f_{ij}(x_i, x_j) \\ & + \sum_{\substack{i,j,k=1 \\ i \neq j \neq k}}^v f_{ijk}(x_i, x_j, x_k) + \dots \end{aligned} \quad (2.5)$$

Persamaan (2.5) menunjukkan bahwa penjumlahan pertama meliputi semua fungsi basis untuk satu variabel. Penjumlahan kedua meliputi semua fungsi basis untuk interaksi antara dua variabel. Penjumlahan ketiga meliputi semua fungsi basis untuk interaksi antara tiga variabel, dan seterusnya. Untuk mempermudah interpretasi model MARS, maka persamaan model MARS pada (2.3) dapat disederhanakan sebagai berikut.

$$\hat{f}(\mathbf{x}) = a_0 + a_1 BF_1 + a_2 BF_2 + a_3 BF_3 + \dots + a_M BF_M \quad (2.6)$$

dengan $\hat{f}(\mathbf{x})$ merupakan variabel respon, a_0 adalah konstanta, a_M adalah koefisien untuk basis fungsi ke- M , dan BF_M adalah basis fungsi ke- M yang didapatkan dari perhitungan rumus

$$B_m(\mathbf{x}) = \prod_{k=1}^{K_m} [s_{km} (x_{v(k,m)} - t_{km})]_+, \text{ serta fungsi basis (BF) yang masuk}$$

pada model didapatkan dari algoritma *backward stepwise*.

Hal-hal yang perlu diperhatikan dalam pembentukan model MARS adalah sebagai berikut (Nash & Bradford, 2001).

1. Knot merupakan akhir dari sebuah garis regresi (*region*) dan awal dari garis regresi (*region*) yang lain. Pada masing-masing titik knot diharapkan adanya kontinuitas dari fungsi basis antar satu *region* dengan *region* lainnya, dengan jarak

antar knot minimum atau observasi antar knot minimum (MO) sebesar 0,1,2, dan 3.

2. Fungsi basis merupakan kumpulan dari fungsi yang digunakan untuk menjelaskan hubungan antara variabel respon dengan variabel prediktor. Fungsi basis terdiri dari satu atau lebih variabel, dan merupakan fungsi parametrik yang didefinisikan pada setiap *region* dengan pada umumnya fungsi basis yang dipilih adalah bentuk polinomial dengan turunan kontinu pada setiap titik knot.
3. Interaksi yaitu *cross product* antar variabel yang saling berhubungan atau berkorelasi. Jumlah interaksi maksimum (MI) yang diperbolehkan adalah 1,2,dan 3. Jika lebih dari 3 maka model akan sulit diinterpretasi.

Langkah awal dalam pembentukan model MARS yaitu menentukan titik-titik perubahan pola perilaku data atau yang disebut dengan titik knot. Pemilihan *knot* pada MARS menggunakan algoritma *forward stepwise* dan *backward stepwise* yang didasarkan pada nilai *Generalized Cross Validation* (GCV) terkecil (Friedman J. H., 1990). Tahap *forward* dilakukan untuk mendapatkan jumlah basis fungsi maksimum dengan kriteria pemilihan basis fungsi adalah meminimumkan *average sum of square residual* (ASR). Kemudian untuk memenuhi konsep parsemoni (model sederhana) dilakukan dengan tahap *backward* yaitu dengan memilih basis fungsi yang dihasilkan dari tahap *forward* dengan meminimumkan nilai *Generalized Cross Validation* atau GCV (Friedman & Silverman, 1989). Bentuk GCV minimum sebagai kriteria untuk menentukan knots ditulis dalam persamaan berikut (Friedman & Silverman, 1989).

$$GCV(M) = \frac{\frac{1}{N} \sum_{i=1}^N [y_i - \hat{f}_M(x_i)]^2}{\left[1 - \frac{\hat{C}(M)}{N}\right]^2} \quad (2.7)$$

dengan,

N : jumlah data

M : jumlah fungsi basis pada model MARS

$\hat{C}(M): C(M) + d.M$

d : nilai ketika fungsi basis mencapai optimasi dari kelipatan 2 hingga 4 kali dari jumlah variabel prediktor ($2 \leq d \leq 4$)

$C(M)$: $\text{trace} [\mathbf{B}(\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T] + 1$

y_i : nilai variabel respon

$\hat{f}_M(x_i)$: nilai taksiran variabel respon pada fungsi basis M

Untuk mendapat model terbaik dalam pemodelan MARS dapat digunakan kriteria nilai GCV paling minimum. Jika terdapat model yang mempunyai nilai GCV yang sama, maka dapat digunakan kriteria MSE yang paling minimum, serta nilai R^2 dan nilai ketepatan klasifikasi yang paling tinggi (Eubank, 1999).

$$MSE = \frac{SSE}{n - p - 1} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - p - 1} \text{ dan } R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.8)$$

Pada model MARS yang dihasilkan dilakukan pengujian koefisien Fungsi Basis (BF) yang meliputi uji serentak dan uji individu. Pengujian koefisien yang dilakukan secara bersamaan atau serentak terhadap fungsi basis yang ada dalam model MARS. Pengujian bertujuan untuk mengetahui apakah secara umum model MARS yang terpilih merupakan model yang sesuai dan menunjukkan hubungan yang tepat antara variabel prediktor dengan variabel respon, berikut hipotesis yang digunakan.

$$H_0 : a_1 = a_2 = a_3 = \dots = a_M = 0$$

$$H_1 : \text{minimal terdapat satu } a_j \neq 0, j = 1, 2, \dots, M$$

$$\text{Statistik uji : } F_{hitung} = \frac{MSR}{MSE} = \frac{((SSR)/p)}{((SSE)/(n-p-1))} \quad (2.9)$$

$$\text{Dengan } SSE = \sum_{i=1}^n (y_i - \hat{y})^2 \text{ dan } SSR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2, \text{ dan hipo-}$$

tesis H_0 ditolak apabila nilai F_{hitung} lebih dari F_{tabel} , dengan nilai derajat bebas (df) = p dan n-p-1, serta nilai tingkat signifikansi sebesar α . Sedangkan pengujian parsial (individu) digunakan untuk mengetahui apakah setiap variabel prediktor mempunyai

pengaruh signifikan terhadap variabel respon pada fungsi basis yang terbentuk didalam model serta mampu untuk menggambarkan keadaan data yang sebenarnya, dengan hipotesis sebagai berikut.

$$H_0 : a_j = 0$$

$$H_1 : a_j \neq 0, j = 1, 2, \dots, M$$

$$\text{Statistik uji : } |T_{hitung}| = \frac{\hat{B}_j}{SE(\hat{B}_j)} \quad (2.10)$$

Nilai $|T_{hitung}|$ yang didapatkan dibandingkan dengan nilai tabel T dengan derajat bebas (df) = n-k dan tingkat signifikansi sebesar α . Apabila nilai $|T_{hitung}|$ bernilai lebih besar T_{tabel} maka didapatkan keputusan Tolak H_0 (Utami, 2016).

2.2 Klasifikasi MARS Respon Biner

Klasifikasi pada model MARS didasarkan pada pendekatan analisis regresi. Klasifikasi pada variabel respon terdiri dari dua nilai sehingga disebut regresi dengan *binary response* (Cox & Snell, 1989). Model probabilitas yang digunakan untuk klasifikasi ditunjukkan pada persamaan berikut:

$$\hat{\pi}(\mathbf{x}) = \frac{e^{\hat{f}(x)}}{1 + e^{\hat{f}(x)}} \text{ dan } 1 - \hat{\pi}(\mathbf{x}) = \frac{1}{1 + e^{\hat{f}(x)}} \quad (2.11)$$

dengan, $\hat{f}(x)$: logit $\hat{\pi}(x)$, probabilitas ($Y=0$) adalah $\hat{\pi}(x)$ dan probabilitas ($Y=1$) adalah $1 - \hat{\pi}(x)$. Y merupakan variabel respon biner (dengan kode 0 dan 1 atau kode 1 dan 2) dan m merupakan banyaknya variabel prediktor, $\mathbf{x} = (x_1, \dots, x_m)$, maka model MARS respon biner untuk klasifikasi dapat dinyatakan sebagai berikut:

$$\begin{aligned} \text{logit } \hat{\pi}(\mathbf{x}) &= \ln \left(\frac{\hat{\pi}(\mathbf{x})}{1 - \hat{\pi}(\mathbf{x})} \right) \\ &= \alpha_0 + \sum_{m=1}^M \alpha_m \prod_{k=1}^{Km} [s_{km}(x_{v(k,m)} - t_{tm})] \end{aligned} \quad (2.12)$$

Persamaan tersebut jika dituliskan dalam bentuk matriks adalah sebagai berikut:

logit $\hat{\pi}(\mathbf{x}) = \mathbf{B}\boldsymbol{\alpha}$, dengan \mathbf{B} terdapat pada persamaan (2.2).

Pengklasifikasian pada model MARS respon biner dilakukan untuk melihat seberapa besar ketepatan dalam mengelompokkan sekumpulan data untuk digolongkan dengan tepat pada kelompoknya. Menurut Agresti (1990), metode klasifikasi yang baik akan menghasilkan kesalahan klasifikasi kecil atau peluang kesalahan klasifikasi (alokasi) yang kecil.

2.3 Ketepatan dan Pengujian Kestabilan Klasifikasi

Total Accuracy Rate (TAR) digunakan untuk menghitung ketepatan klasifikasi pada hasil pengelompokan. Nilai TAR dapat menyatakan representasi proporsi sampel yang tepat diklasifikasikan. Penentuan ketepatan klasifikasi MARS respon biner dengan perhitungan pada tabel klasifikasi berikut.

Tabel 2.1 Klasifikasi MARS Respon Biner

Hasil Observasi	Taksiran Observasi	
	y_0	y_1
y_0	n_{00}	n_{01}
y_1	n_{10}	n_{11}

Keterangan:

y_0 : Tidak ada kejadian Infeksi Oportunistik

y_1 : Ada kejadian Infeksi Oportunistik

n : jumlah observasi

n_{00} : jumlah observasi dari y_0 yang tepat diklasifikasi sebagai y_0

n_{11} : jumlah observasi dari y_1 yang tepat diklasifikasi sebagai y_1

n_{01} : jumlah observasi dari y_0 yang salah diklasifikasi sebagai y_1

n_{10} : jumlah observasi dari y_1 yang salah diklasifikasi sebagai y_0

Nilai *Total Accuracy Rate* (TAR) didapatkan dengan perhitungan sebagai berikut:

$$TAR(\%) = \frac{\text{jumlah prediksi benar}}{\text{jumlah total prediksi}} = \frac{n_{00} + n_{11}}{n} \times 100\% \quad (2.13)$$

Sedangkan nilai APER yang digunakan untuk menunjukkan besarnya kesalahan klasifikasi adalah sebagai berikut:

$$APER(\%) = \frac{\text{jumlah prediksi salah}}{\text{jumlah total prediksi}} = \frac{n_{01} + n_{10}}{n} \times 100\% \quad (2.14)$$

Dalam mengukur klasifikasi dipertimbangkan pula nilai *sensitivity* yang menggambarkan akurasi pada sampel kelas positif. *Specificity* untuk menggambarkan akurasi pada sampel kelas negatif. Nilai *G-means* mampu menggambarkan sebuah metode klasifikasi melalui *sensitivity* dan *specificity* secara bersamaan. semakin besar nilai *G-means* menunjukkan bahwa metode klasifikasi tersebut mampu memprediksi data ditiap kelas dengan baik dan sesuai digunakan untuk data yang tidak seimbang. Metode klasifikasi yang baik harusnya yang mampu mengukur *sensitivity*, *specificity* dan *G-means* yang sama baiknya.

$$\begin{aligned} \text{Sensitivity} &= \frac{n_{00}}{n_{00} + n_{01}} \\ \text{Specificity} &= \frac{n_{11}}{n_{10} + n_{11}} \end{aligned} \quad (2.15)$$

$$G - \text{means} = \sqrt{\text{sensitivity} \times \text{specificity}}$$

Untuk membandingkan antara jumlah ketepatan klasifikasi dengan total sampel dan jumlah kelompok, maka dapat dilakukan pengujian dengan Hipotesis berikut ini.

H_0 : Klasifikasi belum dianggap stabil dan konsisten secara statistik

H_1 : Klasifikasi dianggap sudah stabil dan konsisten secara statistik

Statistik uji *Press's Q* didapatkan dengan persamaan berikut:

$$\text{Press's } Q = \frac{[N - (nK)]^2}{N(K - 1)} \quad (2.16)$$

dengan,

N : jumlah total sampel

n : jumlah individu yang tepat diklasifikasikan

K : jumlah kelompok

Nilai *Press's Q* dibandingkan dengan nilai kritis pada tabel *chisquare* derajat bebas 1. Jika nilai *Press's Q* melebihi nilai

kritis, maka diambil keputusan Tolak H_0 yang artinya klasifikasi dapat dianggap sudah stabil dan konsisten secara statistik (Hair JR., dkk., 2010).

2.4 *Bootstrap Aggregating (Bagging)*

Bootstrap Aggregating (Bagging) merupakan teknik yang diusulkan oleh Breiman (1996) yang dapat digunakan untuk mereduksi variansi estimator pada metode klasifikasi dan regresi. Penggunaannya tidak dibatasi hanya untuk memperbaiki estimator. Teknik ini juga dapat memperbaiki stabilitas, meningkatkan akurasi, dan kekuatan prediktif. Jumlah replikasi minimum untuk tujuan klasifikasi yakni sebanyak 50 kali, serta optimum saat nilai akurasi tertinggi telah didapatkan (Breiman, 1994). Bagging merupakan salah satu prosedur intensif perhitungan untuk memperbaiki estimator atau pengklasifikasi yang tidak stabil, khususnya masalah dimensi tinggi. Bagging merupakan salah satu bagian dari *Bootstrap*. Perbedaan pada Bagging dilakukan dengan membandingkan sampel secara berpasangan.

Buhlmann dan Yu (2002), secara singkat menyatakan algoritma Bagging sebagai berikut.

1. Sebuah data set L yang terdiri dari $\{(y_i, x_i), i = 1, 2, \dots, n\}$. Melakukan replikasi *bootstrap* pada data, sehingga didapatkan $L_i^* = (y_i^*, x_i^*)$, dengan $i = 1, 2, \dots, n$.
2. Replikasi *bootstrap* dilakukan sebanyak B kali, sehingga didapatkan $(L^{(B)})$ yang merupakan *resampling* pengembalian dari L .

Metode Bagging memiliki potensi untuk menurunkan kuadrat *error* peramalan pada berbagai proses, dan algoritma Bagging dalam pemodelan MARS adalah sebagai berikut (Buhlmann & Yu, 2002).

1. Terdapat data set L yang terdiri dari $\{(y_i, x_i), i = 1, 2, \dots, n\}$, dan melakukan replikasi *bootstrap* pada data, sehingga didapatkan $L_i^* = (y_i^*, x_i^*)$ atau disebut $(L^{(B)})$.
2. Melakukan pemodelan MARS pada $(L^{(B)})$.

3. Memprediksi variabel respon dari model MARS yang telah dihasilkan.
4. Mengulangi langkah 1 sampai dengan 3 hingga B replikasi *bootstrap*.
5. Melakukan prediksi pada variabel respon berdasarkan pemilihan prediksi yang sering muncul pada masing-masing observasi dari B replikasi *bootstrap* (maksimum *voting*).
6. Menghitung ketepatan klasifikasi prediksi model Bagging MARS.

2.5 Pengertian HIV/AIDS

HIV (*Human Immunodeficiency Virus*) merupakan virus yang menyerang sel darah putih, sedangkan penyakit AIDS (*Acquired Immune Deficiency Syndrome*) adalah kumpulan gejala penyakit akibat menurunnya sistem kekebalan tubuh yang disebabkan oleh virus HIV. Penyakit HIV maupun AIDS tergolong dalam suatu kumpulan gejala penyakit (*syndrome*) yang disebabkan oleh *Retrovirus* yang menyerang sistem kekebalan atau pertahanan tubuh. Apabila seseorang telah terinfeksi maka akan mudah diserang penyakit-penyakit lain yang dapat berakibat fatal atau disebut Infeksi Oportunistik. Munculnya *syndrome* tersebut tidak terjadi seketika melainkan membutuhkan waktu sekitar 5 hingga 10 tahun setelah seseorang terinfeksi HIV (Musoke, 2007). Kasus AIDS pertama kali ditemukan oleh Gottlieb di Amerika Serikat pada tahun 1981, sedangkan virusnya ditemukan oleh Luc Montagnier pada tahun 1983 (Sudoyo, dkk., 2007).

Virus HIV sangat lemah dan dapat mati diluar tubuh, sehingga *verikulum* (pembawa *agent*) yang dapat membawa virus HIV keluar tubuh dan menularkan kepada orang lain melalui berbagai cairan tubuh seperti cairan vagina atau serviks, semen, dan darah penderita. WHO telah mengklasifikasikan stadium penyakit HIV/AIDS dari level yang paling permulaan (awal) pada level I sampai terminal penyakit yakni pada level IV. Sebagian besar HIV/AIDS berakibat fatal karena sekitar 75% pasien yang terdiagnosis AIDS akan meninggal.

2.6 Faktor yang Mempengaruhi Kejadian Infeksi Oportunistik Penderita HIV/AIDS

Infeksi Oportunistik adalah infeksi yang timbul akibat menurunnya kekebalan tubuh karena mikroba (bakteri, jamur, dan virus) jika dalam keadaan normal dapat terkendali oleh kekebalan tubuh, sedangkan pada ODHA dapat menyebabkan kematian. Kerusakan pada sistem kekebalan tubuh dapat dicegah dan dihindari dengan penggunaan terapi antiretroviral (ARV) jumlah Infeksi Oportunistik >2 memiliki risiko kematian 2 kali lebih besar dibandingkan dibawahnya (Spiritia, 1995).

Secara garis besar terdapat dua faktor yang membuat menurunnya daya tahan tubuh akibat virus HIV (Infeksi Oportunistik) yakni faktor predisposisi, dan faktor manifestasi klinis. Karakteristik predisposisi adalah faktor-faktor individu yang meliputi demografis (umur, jenis kelamin, status perkawinan), struktur sosial (pendidikan, pekerjaan, jumlah keluarga, ras, suku, agama, perpindahan tempat tinggal), dan keyakinan (penilaian terhadap status sehat sakit, sikap terhadap pelayanan, dan dukungan keluarga) (Nasronuidin, 2012).

Umur penderita HIV/AIDS mempengaruhi ketahanan imunitas pasien, semakin tua umurnya maka akan menyebabkan daya tahan tubuh semakin melemah dan menurunnya fungsi kekebalan tubuh, dengan proporsi terbanyak yakni pasien berjenis kelamin laki-laki. Dari hasil Risksesdas (2010) diketahui bahwa semakin tinggi pendidikan seseorang maka pengetahuan cenderung semakin meningkat dan mempunyai upaya yang lebih tinggi untuk pencegahan dan penanggulangan penyakit HIV/AIDS (Departemen Kesehatan RI, 2010).

Untuk faktor manifestasi klinis meliputi Kadar CD4, stadium klinis, dan terapi ARV. Kesehatan sistem kekebalan tubuh dapat dinilai dengan mengukur jumlah sel CD4. Pada orang yang tidak terinfeksi HIV, jumlah sel CD4 berkisar antara 500 hingga 1.500, setelah terinfeksi HIV/AIDS maka sel CD4 menurun hingga berkisar 200 sel (Spiritia, 1995).

Pengobatan dengan *antiretroviral therapy* (ARV) dimulai pada penderita dengan syarat kadar CD4<350 dan stadium pasien HIV/AIDS. ARV menjadi terapi pilihan bagi pasien karena obat ini dapat memperlambat progresivitas penyakit dan memperbaiki daya tahan tubuh, karena setelah terjadi infeksi HIV maka akan menurunkan sel CD4 secara bertahap dan turunya sel CD4 sejalan dengan meningkatnya stadium sehingga menyebabkan peningkatan gangguan imunitas dan semakin memperpendek kelangsungan hidup pasien (Nasronudin, 2012).

ARV yang tersedia di Indonesia umumnya harus dipakai dua kali sehari dengan kepatuhan tinggi. Virus HIV tetap ada ditubuh sehingga jika tanpa pengobatan akan menggandakan diri, oleh karena itu ARV harus dipakai seumur hidup. Dukungan dan perawatan medis yang baik dari tahap awal penyakit akan lebih berhasil menangani infeksi (Spiritia, 1995).

Berdasarkan analisis penyebab masalah kesehatan, Teori Green membedakan determinan masalah kesehatan menjadi tiga faktor diantaranya adalah:

a. Faktor predisposisi (*predisposing factors*)

Faktor-faktor yang dapat mempermudah atau mengpres-deposisikan terjadinya perilaku pada diri seseorang atau masyarakat adalah pengetahuan dan sikap seseorang atau masyarakat tersebut terhadap apa yang dilakukan.

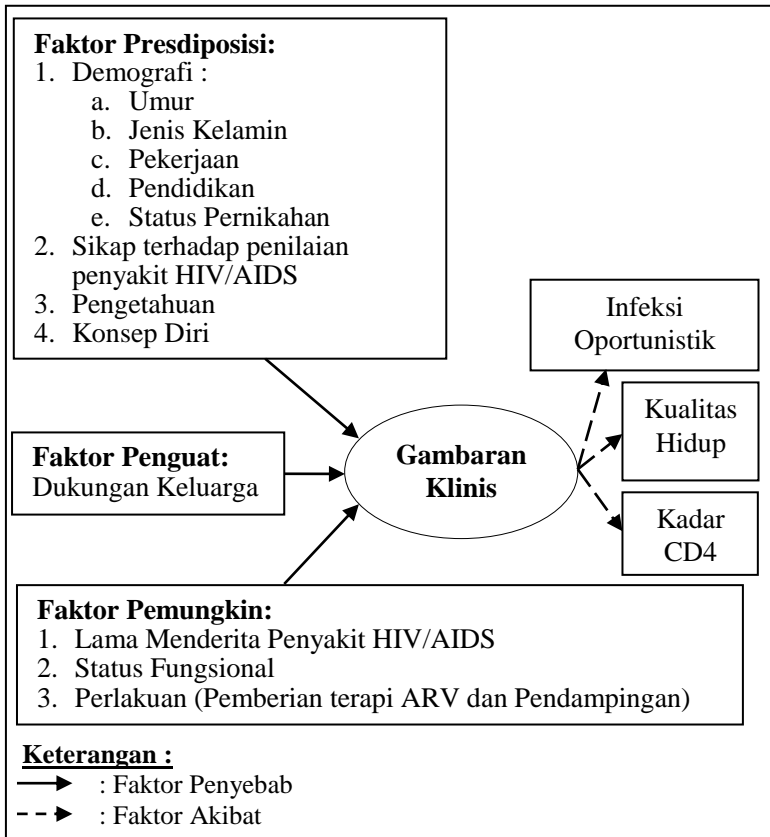
b. Faktor pemungkin (*enabling factors*)

Faktor pemungkin atau pendukung (*enabling*) perilaku adalah fasilitas, sarana, atau prasarana yang mendukung atau yang memfasilitasi terjadinya perilaku seseorang atau masyarakat. Pengetahuan dan sikap saja belum menjamin terjadinya perilaku, maka masih diperlukan sarana atau fasilitas untuk memungkinkan atau mendukung perilaku tersebut.

c. Faktor penguat (*reinforcing factors*)

Pengetahuan, sikap, dan fasilitas yang tersedia kadang-kadang belum menjamin terjadinya perilaku seseorang atau masyarakat, sehingga adanya faktor penguat sebagai faktor utama determinan penyakit.

Untuk menggambarkan penjelasan faktor-faktor yang mempengaruhi kejadian Infeksi Oportunistik dan determinan perilaku penyakit HIV/AIDS lebih lanjut akan dijelaskan pada Gambar 2.1 mengenai kerangka konsep berdasarkan Teori Lawrence Green.



Gambar 2.1 Kerangka Konsep Teori Lawrence Perilaku Penyakit HIV/ AIDS

(Halaman ini sengaja dikosongkan)

BAB III METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan pada penelitian ini merupakan data sekunder yang berasal dari rekam medis pasien HIV/AIDS pada tahun 2016 di Puskesmas Kabupaten Pasuruan sejumlah 92 data yang terisi penuh pada masing-masing variabel yang diteliti. Kriteria pasien yang diteliti adalah yang telah menjalani pengobatan HIV/AIDS lebih dari satu tahun di Puskesmas.

3.2 Variabel Penelitian

Terdapat dua jenis variabel yang digunakan dalam penelitian ini yakni variabel respon dan variabel prediktor untuk menggambarkan model dan peluang kejadian Infeksi Oportunistik pada pasien penderita penyakit HIV/AIDS yang didapatkan dari klasifikasi metode MARS dan Bagging MARS. Penjelasan dari masing-masing variabel ditunjukkan pada Tabel 3.1.

Tabel 3.1 Variabel Penelitian

Variabel	Keterangan	Kategori	Skala Pengukuran
Variabel Respon			
Y	Infeksi Oportunistik (IO)	1. Tidak ada IO (0) 2. Ada IO (1)	Nominal
Variabel Prediktor			
X ₁	Umur	-	Rasio
X ₂	Jenis Kelamin	1. Laki-laki (1) 2. Perempuan (2)	Nominal
X ₃	Pekerjaan	1. PNS (1) 2. Pensiunan (2) 3. Petani (3) 4. Swasta (4) 5. Tidak bekerja (5) 6. Dan lain-lain (6)	Nominal
X ₄	Pendidikan	1. Tidak Sekolah (1) 2. SD/ sederajat (2) 3. SMP/ sederajat (3)	Ordinal

Tabel 3.1 (Lanjutan) Variabel Penelitian

Variabel	Keterangan	Kategori	Skala Pengukuran
X ₄	Pendidikan	4. SMA/ sederajat (4) 5. Akademi DI hingga DIII (5) 6. Perguruan Tinggi S1 hingga S3 (6)	Ordinal
X ₅	Status Pernikahan	1. Kawin (1) 2. Tidak kawin (2) 3. Cerai (3) 4. Tidak terbuka (4)	Nominal
X ₆	Lama terinfeksi virus HIV	-	Rasio
X ₇	Kadar CD4 Akhir	-	Rasio
X ₈	Perlakuan	1. Hanya mendapat ARV (0) 2. Mendapat ARV dan pendampingan (1)	Nominal
X ₉	Pengetahuan	1. Baik (0) 2. Kurang baik (1)	Nominal
X ₁₀	Konsep Diri	1. Positif (0) 2. Negatif (1)	Nominal
X ₁₁	Dukungan Keluarga	1. Mendukung (0) 2. Tidak mendukung (1)	Nominal
X ₁₂	Kualitas Hidup	1. Baik (0) 2. Kurang baik (1)	Nominal

Penjelasan untuk beberapa variabel penelitian yang digunakan adalah sebagai berikut.

1. Infeksi Oportunistik (Y): Merupakan jenis gangguan penyerta (penyakit infeksi selain HIV) pada penderita penyakit HIV/AIDS seperti TB paru, hepatitis, sipilis, dan lain sebagainya berdasarkan diagnosa dokter. Jika ada satu saja penyakit infeksi maka termasuk ada IO.
2. Kadar CD4 akhir (X₇): Menyatakan perentase sel T-CD4 yang diserang virus HIV dibandingkan dengan total sel T-

CD4 yang ada. Kadar CD4 ini diambil yang terakhir kali saat dilakukan pengobatan ARV.

3. Perlakuan (X_8): Jenis perlakuan dibedakan menjadi 2 yakni hanya dengan terapi ARV (*Antiretroviral Therapy*) dan adanya terapi ARV serta pengobatan HIV/AIDS lainnya pada penderita HIV dan pendampingan dari sebaya.
4. Pengetahuan (X_9): Sesuatu yang diketahui oleh pasien meliputi pengertian, tanda gejala, pengobatan, dan cara mencegah penularan virus HIV.
5. Konsep Diri (X_{10}): Konsep diri pada pasien HIV/AIDS yang lebih menghargai dirinya sendiri karena menganggap dirinya dalam keadaan kurang baik dan klinik VCT membantu memperbaiki kesehatan dan kepribadian.
6. Dukungan Keluarga (X_{11}): Persepsi pasien terhadap dukungan yang diberikan oleh keluarga meliputi dukungan emosional, informasi, sosial dan lain sebagainya.
7. Kualitas Hidup (X_{12}): Persepsi pasien terhadap kualitas hidup yang mencakup kesejahteraan fisik, psikologis, hubungan sosial, dan kesejahteraan spiritual.

3.3 Metode Pengolahan data

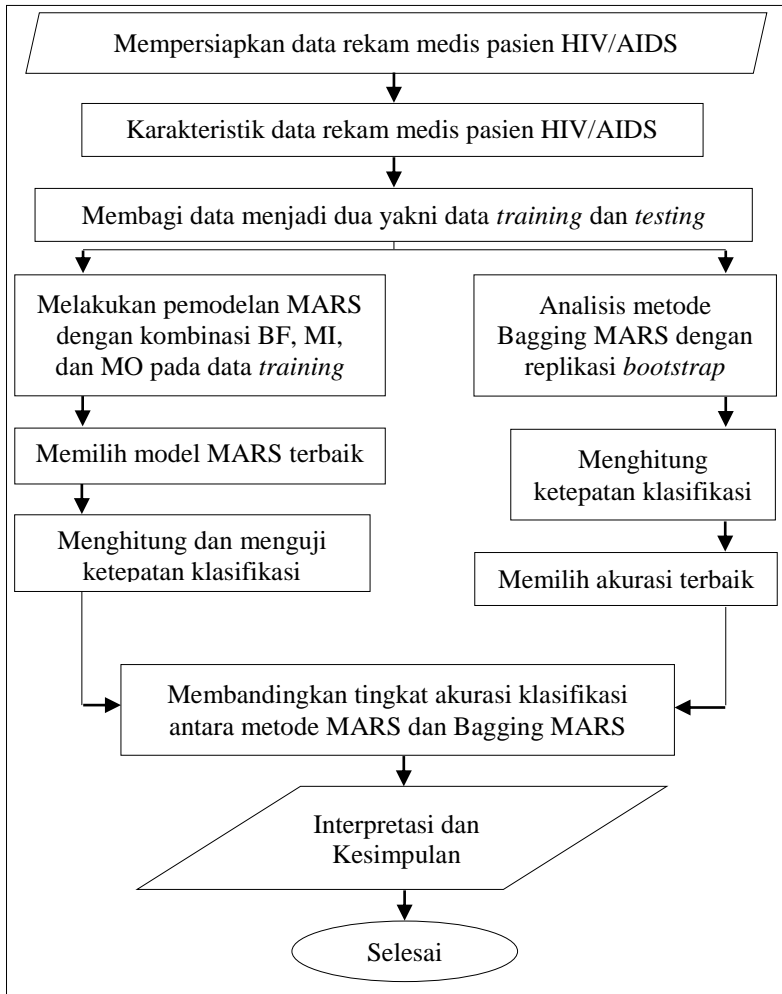
Pengolahan data diawali dengan penyortiran data yang akan diolah, dengan tidak mengikutsertakan data yang tidak terdapat variabel respon dan prediktor (tidak terisi). Berikut langkah analisis yang digunakan dalam melakukan penelitian ini.

1. Untuk menjawab tujuan pertama dilakukan analisis statistika deskriptif untuk mendeskripsikan karakteristik faktor-faktor yang diduga mempengaruhi Infeksi Oportunistik pada pasien HIV/AIDS.
2. Untuk menjawab tujuan kedua dilakukan langkah analisis sebagai berikut:
 - a. Membagi data menjadi dua yakni data *training* untuk membangun model dan data *testing* untuk validasi data serta melihat kemampuan model untuk memprediksi.

- b. Melakukan pemodelan MARS pada data rekam medis pasien HIV/AIDS dengan mengkombinasikan besarnya *Basis Function* (BF), *Maximum Interaction* (MI), dan *Minimum Observation* (MO).
 - c. Menentukan model terbaik dari MARS berdasarkan nilai GCV terkecil.
 - d. Mengelompokkan fungsi basis berdasarkan variabel prediktor yang masuk dalam model.
 - e. Menginterpretasikan tingkat kepentingan variabel dan pengurangan GCV variabel.
 - f. Menghitung nilai ketepatan klasifikasi dan menguji kestabilan klasifikasi berdasarkan model MARS terbaik.
 - g. Melakukan pemodelan Bagging MARS pada data dengan replikasi 50, 100, 150, 200, dan 500 kali.
 - h. Memilih akurasi ketepatan klasifikasi terbaik dari hasil replikasi metode Bagging MARS berdasarkan nilai persentase keakuratan tertinggi
3. Untuk menjawab tujuan ketiga dilakukan dengan membandingkan tingkat akurasi klasifikasi antara metode MARS dan Bagging MARS untuk mendapatkan metode terbaik.
 4. Menginterpretasi hasil analisis serta memberikan kesimpulan dan saran.

3.4 Diagram Alir (*Flowchart*)

Berikut ini adalah diagram alir yang menggambarkan langkah analisis data dalam analisis dan pembahasan pada Tugas Akhir ini.



Gambar 3. 1 Diagram Alir Pemodelan MARS dan Bagging MARS

(Halaman ini sengaja dikosongkan)

BAB IV

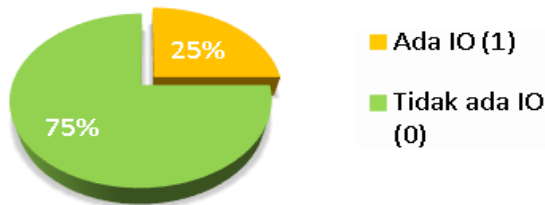
ANALISIS DAN PEMBAHASAN

Dalam analisa dan pembahasan kali ini digunakan data mengenai klasifikasi kejadian Infeksi Oportunistik pada pasien HIV/AIDS di Kabupaten Pasuruan dengan unit analisis pada penelitian ini adalah individu yang menderita penyakit HIV/AIDS dan menjalani pengobatan di Puskesmas. Metode yang digunakan dalam analisis dan pembahasan kali ini yaitu dengan *Multivariate Adaptive Regression Splines* (MARS) dan *Bootstrap Aggregating* MARS (Bagging MARS). Model pengklasifikasian kejadian Infeksi Oportunistik pada pasien HIV/AIDS didapatkan melalui prediksi faktor-faktor yang mempengaruhi serta didapatkan peluang kejadian dan tingkat akurasi klasifikasi kasus Infeksi Oportunistik pasien yang terkena HIV/AIDS. Namun sebelumnya dilakukan analisa deskriptif untuk mendeskripsikan karakteristik dari data dan menunjukkan perbandingan pembagian data *training* dan *testing* yang tepat sehingga didapatkan hasil yang akurat untuk mengklasifikasi.

4.1 Karakteristik Pasien Penderita Penyakit HIV/AIDS

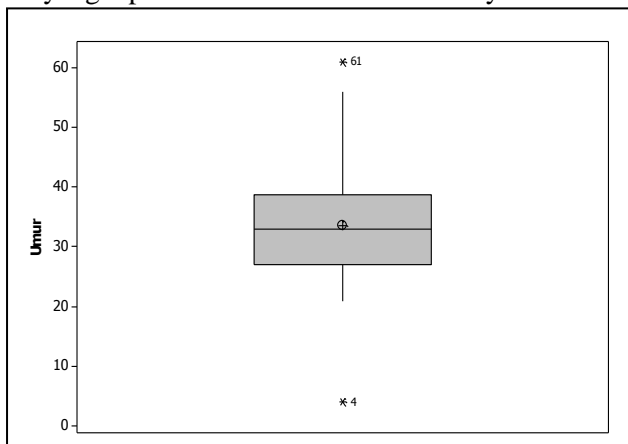
Uraian untuk mengetahui informasi yang dapat diperoleh dari pasien penderita penyakit HIV/AIDS dapat diketahui melalui analisis statistika deskriptif yang digambarkan dengan grafik *piechart*, *barchart* untuk data kategorik, serta *boxplot* untuk data kontinu sehingga dapat menggambarkan nilai rata-rata, median, sebaran data, dan *outlier* dari data.

Penderita HIV/AIDS mempunyai kendala utama pada adanya Infeksi Oportunistik (IO) yakni infeksi akibat menurunnya kekebalan tubuh karena mikroba (bakteri, jamur, dan virus). Jika dalam keadaan normal dapat terkendali oleh kekebalan tubuh, sedangkan pada ODHA dapat menyebabkan kematian jika tidak dilakukan pengobatan, terapi, serta beberapa penanganan khusus lainnya. Berikut merupakan gambaran kejadian Infeksi Oportunistik (IO) pada pasien.



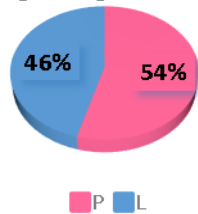
Gambar 4.1 Persentase Kejadian Infeksi Oportunistik Pasien HIV/AIDS

Dapat diketahui pada Gambar 4.1 yang menunjukkan persentase kejadian Infeksi Oportunistik (IO) pasien HIV/AIDS di Puskesmas berdasarkan diagnosa dokter. Jika ada satu saja penyakit infeksi seperti TB paru, hepatitis, sipilis, dan lain sebagainya maka dikategorikan menjadi ada IO (kategori 1) didalam tubuh pasien, sedangkan jika tidak terdapat infeksi maka dikategorikan tidak ada IO (kategori 0) pada pasien. Dari hasil klasifikasi menunjukkan bahwa 75% atau sejumlah 69 orang pasien tidak ditemukan adanya IO dan hanya 25% atau sejumlah 23 orang yang mengalami kejadian IO. Adanya perbedaan hasil tersebut diharapkan dapat diketahui faktor-faktor apa yang menyebabkan pasien terkena infeksi penyakit dan hasil akurasi klasifikasi yang tepat berdasarkan faktor-faktornya.



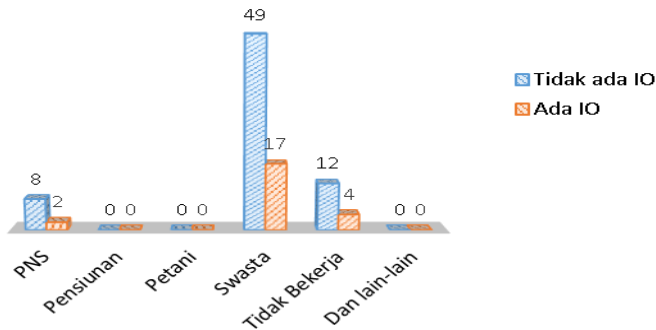
Gambar 4.2 Boxplot Umur Pasien HIV/AIDS

Variabel umur digunakan dalam prediktor klasifikasi kejadian IO karena semakin tua umur seorang pasien HIV/AIDS menyebabkan daya tahan tubuh semakin melemah dan menurunnya fungsi kekebalan tubuh. Gambar 4.2 menunjukkan bahwa rata-rata umur pasien HIV/AIDS adalah berusia 33 tahun, dengan umur terendah pasien yang terkena penyakit HIV/AIDS adalah berusia 4 tahun dan yang tertinggi adalah pasien yang berusia 61 tahun dan keduanya merupakan pencilan dari data atau *outlier*.



Gambar 4.3 Persentase Jenis Kelamin Pasien HIV/AIDS

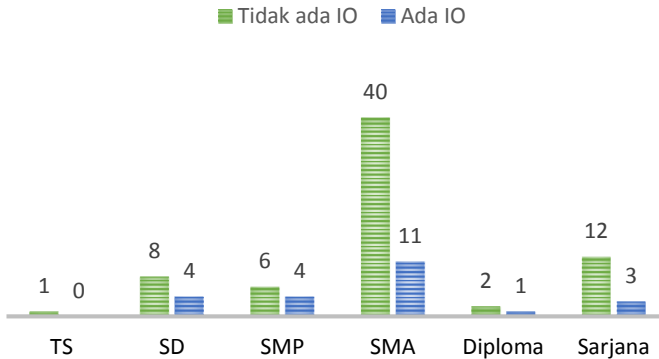
Gambar 4.3 menunjukkan karakteristik pasien berdasarkan jenis kelamin dan menunjukkan bahwa terdapat jumlah data yang hampir seimbang antara pasien perempuan maupun pasien laki-laki dengan masing-masing sebesar 54% untuk pasien HIV/AIDS perempuan dan 46% untuk pasien HIV/AIDS laki-laki.



Gambar 4.4 Jumlah Pasien HIV/AIDS Berdasarkan Kriteria Pekerjaan dan Kejadian Infeksi Oportunistik

Berdasarkan kriteria pekerjaan pasien yang ditunjukkan pada Gambar 4.4 dapat diketahui bahwa pasien yang bekerja swasta (kategori 4) merupakan yang terbanyak atau sejumlah 66 orang

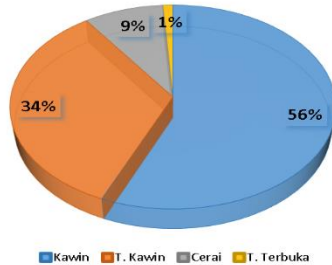
pasien HIV/AIDS dengan 17 orang yang mengalami Infeksi Oportunistik (IO) serta 49 orang pasien yang tidak mengalami IO. Jenis pekerjaan terbanyak selanjutnya yakni yang tidak bekerja (kategori 5) dengan jumlah pasien yang mengalami IO sebanyak 4 orang dan 12 orang tidak ada IO, jenis pekerjaan PNS (kategori 1) mempunyai 2 orang pasien yang mengalami IO dan 8 orang pasien yang tidak mengalami IO. Pada jenis pekerjaan lain seperti pensiunan (kategori 2), petani (kategori 3), dan jenis pekerjaan yang lain (kategori 6) tidak terdapat adanya pasien penderita HIV/AIDS.



Gambar 4.5 Jumlah Pasien HIV/AIDS Berdasarkan Kriteria Pendidikan dan Kejadian Infeksi Oportunistik

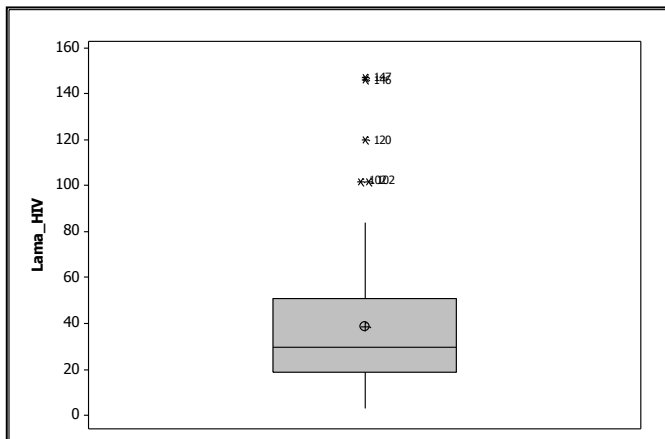
Karakteristik pasien HIV/AIDS berdasarkan pendidikan pada Gambar 4.5 menunjukkan bahwa semua kategori jenjang pendidikan terdapat adanya pasien HIV/AIDS. Pada jenjang pendidikan lulusan SMU/ sederajat (kategori 4) merupakan yang terbanyak yakni sejumlah 51 orang dengan rincian 22% orang atau sejumlah 11 orang yang mengalami IO dan 78% atau sejumlah 40 orang tidak mengalami IO. Pasien dengan pendidikan lulusan Perguruan Tinggi (S1-S3) dengan kategori 6 merupakan kedua terbanyak yakni sejumlah 15 orang dengan 80% atau sejumlah 12 orang tidak mengalami IO dan 20% atau sejumlah 3 orang mengalami IO. Untuk jenjang pendidikan SD/ sederajat (2) menunjukkan bahwa 33% pasien mengalami IO dan 67% tidak

mengalami IO, pada jenjang pendidikan SLTP/ sederajat (3) terdapat 40% orang mengalami IO dan 50% orang pasien yang tidak sekolah mengalami peluang terjadinya IO terbanyak.



Gambar 4.6 Jumlah Pasien HIV/AIDS Berdasarkan Kriteria Status Pernikahan

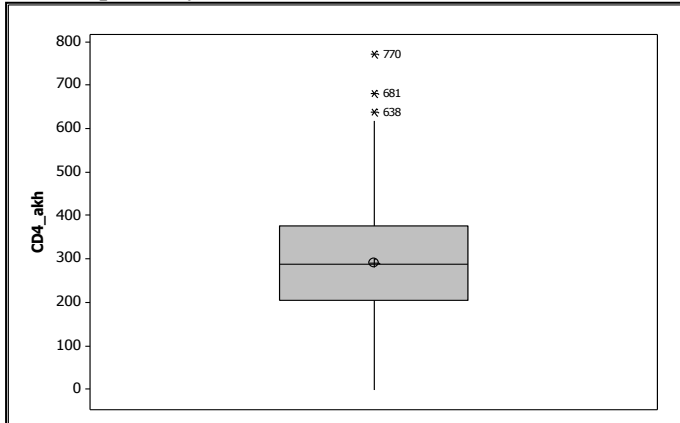
Pada Gambar 4.6 yang menunjukkan status pernikahan pasien HIV/AIDS dapat diketahui bahwa terdapat 56% pasien telah menikah atau kawin, sedangkan pasien yang tidak atau belum melakukan pernikahan terdapat sebesar 34% pasien, 9% pasien yang mengalami perceraian, dan terdapat 1% pasien yang tidak terbuka saat ditanya dan melakukan pengobatan di Puskesmas.



Gambar 4.7 Boxplot Lama Terjadinya Penyakit HIV/AIDS

Rata-rata lama waktu kejadian penyakit HIV/AIDS yang ditunjukkan pada Gambar 4.7 adalah berkisar selama 38,61 bulan

atau kurang lebih selama 3 tahun, dengan lama waktu terjadinya penyakit HIV/AIDS pada pasien paling rendah adalah selama 3 bulan dan yang paling lama yakni selama 147 bulan atau berkisar 12 tahun, serta terdapat data *outlier* pada pasien HIV/AIDS yang sudah cukup lama yakni selama 102, 120, 146, dan 147 bulan.

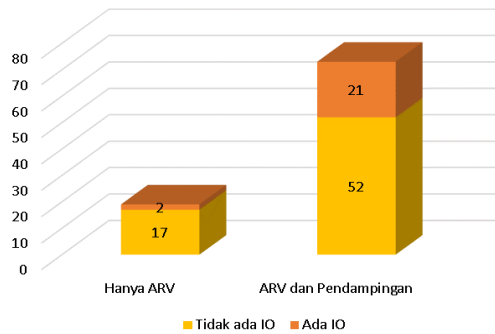


Gambar 4.8 Boxplot Kadar CD4 pada Pasien HIV/AIDS

Virus HIV menyerang sel darah putih penderitanya, sehingga kekebalan tubuhnya akan semakin menurun dan mudah mengalami infeksi jika tidak diberikan pengobatan dan terapi. Kesehatan sistem kekebalan tubuh dapat dinilai dengan mengukur jumlah sel CD4 dalam darah sehingga semakin kecil sel CD4 maka semakin besar peluang terjadinya infeksi. Gambar 4.8 menunjukkan bahwa rata-rata jumlah sel CD4 dalam darah pasien penyakit HIV/AIDS adalah sebesar 291,4 dan yang paling tinggi adalah sejumlah 770, hal ini tepat karena jumlah sel CD4 pada pasien yang telah terinfeksi HIV/AIDS maka sel CD4 menurun hingga berkisar 200 sel. Jumlah CD4 yang tinggi menunjukkan bahwa pasien di Puskesmas sudah diberikan pengobatan, terapi, dan dukungan yang baik.

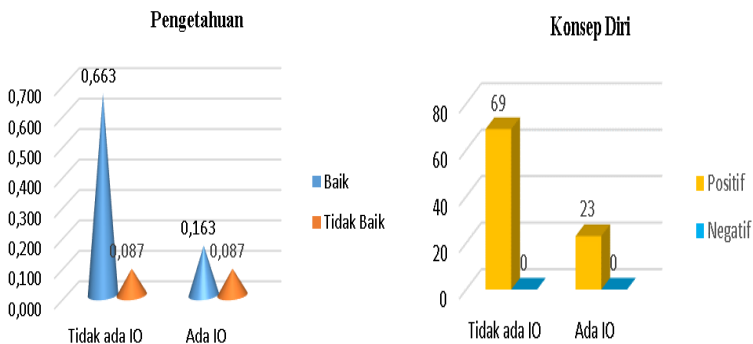
Perlakuan yang diberikan pada pasien HIV/AIDS antarlain adanya terapi ARV (*Antiretroviral Therapy*) dan pengobatan HIV/AIDS lainnya pada pasien serta pendampingan dari sebaya namun masih terdapat pasien yang belum mendapatkan pendam-

pingan dan untuk menunjukkan jumlah pasien yang belum mendapat pendampingan dapat ditunjukkan pada Gambar 4.9.



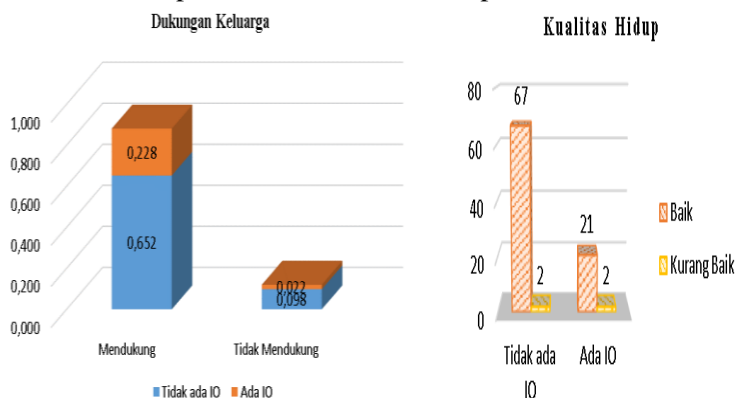
Gambar 4.9 Jumlah Pasien HIV/AIDS Berdasarkan Kriteria Perlakuan dan Kejadian Infeksi Oportunistik

Gambar 4.9 dapat diketahui bahwa terdapat 19 pasien atau berkisar 21% yang hanya mendapat terapi ARV (kategori 0) dengan 89% atau sejumlah 17 orang pasien yang tidak mengalami adanya Infeksi Oportunistik (IO) dan 11% atau sejumlah 2 orang pasien yang mengalami terjadinya kasus IO, sedangkan perlakuan pasien yang mendapat ARV dan pendampingan adalah sebesar 79% atau 73 orang pasien dan diantaranya terdapat 52 orang pasien tidak mengalami terjadinya IO sedangkan 21 orang pasien mengalami IO.



Gambar 4.10 Jumlah Pasien HIV/AIDS Berdasarkan Kriteria Pengetahuan dan Konsep Diri pada Pasien HIV/AIDS

Gambar 4.10 menunjukkan jumlah pasien HIV/AIDS berdasarkan kriteria pengetahuan serta konsep dirinya, dan dapat diketahui bahwa pengetahuan pasien sudah dikatakan baik (kategori 0) karena 82,60% sudah mengetahui adanya tanda gejala, pengobatan dan cara mencegah penularan penyakit HIV/AIDS dan 17,40% masih kurang mengerti (kategori 1) mengenai pengetahuan tersebut. Variabel kosep diri menunjukkan bahwa 100% pasien HIV/AIDS lebih menghargai dirinya sendiri karena menganggap dirinya dalam keadaan kurang baik dan klinik VCT membantu memperbaiki kesehatan dan kepribadian.



Gambar 4.11 Jumlah Pasien HIV/AIDS Berdasarkan Kriteria Dukungan Keluarga dan Kualitas Hidup Pasien HIV/AIDS

Dukungan keluarga yang diberikan pada pasien HIV/AIDS adalah keluarga yang menunjukkan kepedulian serta memperhatikan keadaan pasien yang ditunjukkan pada Gambar 4.11 dapat diketahui bahwa 88% keluarga pasien mendukung (kategori 0) dengan 65,20% pasien tidak mengalami kasus IO serta 22,80% pasien mengalami kasus IO. Sedangkan kategori kualitas hidup pasien menunjukkan bahwa 95,66% kualitas hidupnya baik (kategori 0) atau merasa sejahtera baik fisik, psikologis, hubungan sosial, dan lain sebagainya.

4.2 Klasifikasi dan Pemodelan Kejadian Infeksi Oportunistik pada Pasien HIV/AIDS Menggunakan Metode *Multivariate Adaptive Regression Splines* (MARS)

Untuk menganalisis menggunakan metode MARS dibutuhkan data *training* yang digunakan membangun model dan data *testing* yang digunakan untuk memvalidasi data serta melihat kemampuan model dalam memprediksi. Namun pembagian data *training* dan data *testing* yang tepat digunakan perlu diperhatikan, sehingga Tabel 4.1 menunjukkan perbandingan hasil pembagian data agar analisis selanjutnya lebih tepat menggunakan pembagian data tersebut.

Tabel 4.1 Kriteria Pembagian Data *Training* dan *Testing*

Perbandingan data	Jenis Data	Nilai GCV	MSE	R ²	Ketepatan Klasifikasi
75 – 25	Training (69)	0,151	0,066	0,731	0,928
	Testing (23)	-	-	-	0,478
80 – 20*	Training (73)	0,178	0,083	0,689	0,918
	Testing (19)	-	-	-	0,579
90 – 10*	Training (82)	0,173	0,098	0,636	0,866
	Testing (10)	-	-	-	0,500

*Resume Model Terbaik dari Lampiran 7 dan 8

Untuk mendapatkan pembagian data *training* dan data *testing* yang tepat dapat dilihat berdasarkan nilai GCV, MSE, R², serta keakuratan klasifikasi dari model pada data *training* dan hasilnya menunjukkan bahwa pembagian data 75% (sejumlah 69 data) untuk *training* dan 25% (sejumlah 23 data). Data *testing* menunjukkan perbandingan yang tepat karena memiliki nilai GCV dan MSE terkecil serta nilai R² dan keakuratan klasifikasi yang besar sehingga untuk analisis selanjutnya digunakan pembagian sejumlah data tersebut.

Kemudian dilakukan analisis menggunakan metode MARS dan hal-hal yang perlu diperhatikan dalam pembentukan model MARS adalah fungsi basis (BF) yang merupakan kumpulan dari fungsi yang digunakan untuk menjelaskan hubungan antara variabel respon dengan variabel prediktor dan umumnya sejumlah

dua hingga empat kali jumlah variabel prediktor. Jumlah interaksi maksimum (MI) yang diperbolehkan dan jarak antar knot minimum atau observasi antar knot minimum (MO).

Banyaknya maksimum interaksi (MI) pada model yang digunakan adalah 1, 2, dan 3 karena jika MI lebih dari 3 maka nilai GCV akan semakin meningkat dan model yang dihasilkan akan semakin kompleks (Friedman, 1990). Jika MI yang digunakan sebesar 1 artinya tidak terdapat interaksi antar variabel prediktor pada model, untuk MI sebesar 2 maka artinya bisa terdapat interaksi antar variabel pada model dengan maksimum interaksi adalah 2 variabel, sedangkan untuk MI sebanyak 3 artinya pada model MARS bisa terjadi interaksi maksimal 3 variabel. Untuk minimum observasi (MO) yang digunakan pada penelitian ini adalah 0, 1, 2, dan 3, sehingga langkah selanjutnya adalah melakukan kombinasi BF, MI, dan MO melalui *trial* dan *error* yang kemudian akan dipilih model terbaik berdasarkan nilai GCV terkecil dan nilai R^2 serta ketepatan klasifikasi yang terbesar sebagai berikut.

Tabel 4.2 *Trial and Error* Pembentukan Model MARS untuk Klasifikasi Kejadian Infeksi Oportunistik

Nomor Model	Kombinasi			GCV	MSE	R^2	Keakuratan Klasifikasi
	BF	MI	MO				
1	24	1	0	0,160	0,147	0,345	0,797
2	24	1	1	0,175	0	0,217	0,217
3	24	1	2	0,175	0	0,217	0,217
4	24	1	3	0,160	0,147	0,345	0,797
5	24	2	0	0,172	0,155	0,306	0,812
6	24	2	1	0,163	0,147	0,345	0,826
7	24	2	2	0,164	0,147	0,345	0,797
8	24	2	3	0,163	0,147	0,345	0,797
9	24	3	0	0,172	0,155	0,306	0,812
10	24	3	1	0,151	0,066	0,731	0,928
11	24	3	2	0,157	0,127	0,439	0,841

Tabel 4.2 (Lanjutan) *Trial and Error* Pembentukan Model MARS
untuk Klasifikasi Kejadian Infeksi Oportunistik

Nomor Model	Kombinasi			GCV	MSE	R ²	Keakuratan Klasifikasi
	BF	MI	MO				
12	24	3	3	0,163	0,147	0,345	0,797
13	36	1	0	0,175	0	0,217	0,217
14	36	1	1	0,175	0	0,217	0,217
15	36	1	2	0,175	0	0,217	0,217
16	36	1	3	0,160	0,147	0,345	0,797
17	36	2	0	0,170	0,14	0,383	0,725
18	36	2	1	0,163	0,147	0,345	0,826
19	36	2	2	0,165	0,147	0,345	0,797
20	36	2	3	0,163	0,147	0,345	0,797
21	36	3	0	0,172	0,155	0,306	0,812
22	36	3	1	0,139	0,074	0,693	0,899
23	36	3	2	0,158	0,127	0,439	0,841
24	36	3	3	0,163	0,147	0,345	0,797
25	48	1	0	0,175	0	0,217	0,217
26	48	1	1	0,175	0	0,217	0,217
27	48	1	2	0,175	0	0,217	0,217
28	48	1	3	0,175	0	0,217	0,217
29	48	2	0	0,170	0,14	0,383	0,725
30	48	2	1	0,163	0,147	0,345	0,826
31	48	2	2	0,165	0,147	0,345	0,797
32	48	2	3	0,163	0,147	0,345	0,826
33	48	3	0	0,172	0,155	0,306	0,812
34	48	3	1	0,139	0,074	0,693	0,899
35	48	3	2	0,158	0,127	0,439	0,841
36	48	3	3	0,163	0,147	0,345	0,797

Tabel 4.2 menunjukkan model hasil kombinasi BF, MI, dan MO pada pembentukan model MARS, dengan pemilihan model

terbaik berdasarkan nilai GCV terkecil. Apabila nilai GCV yang didapatkan bernilai sama, maka dapat mempertimbangkan nilai R^2 terbesar apabila masih sama maka dapat digunakan nilai ketepatan klasifikasi terbesar dan jika masih bernilai sama dengan model lain maka pertimbangan selanjutnya dengan memilih model yang memiliki nilai kombinasi BF, MI, dan MO terkecil.

Model MARS terbaik untuk klasifikasi kejadian IO yakni pada urutan model ke-10. Model tersebut dipilih karena meskipun nilai GCV yang dihasilkan lebih besar dibandingkan dari beberapa model lain, namun kriteria yang lain lebih memenuhi untuk nilai MSE, R^2 , dan keakuratan klasifikasi pada model yang dihasilkan jauh lebih baik dengan kombinasi BF sebesar 24, MI sebesar 3, MO sebesar 1, dan nilai GCV sebesar 0,151, MSE terkecil bernilai 0,066, R^2 terbesar bernilai 0,731 serta nilai ketepatan klasifikasi terbesar bernilai 0,928 persen, berikut model yang didapatkan untuk klasifikasi kejadian IO.

$$\hat{f}(x) = 0,993 - 1,000 * BF1 - 0,004 * BF8 + 0,673 * BF11 + 0,031 * BF13 - 0,029 * BF15 + 0,253 * BF17 + 0,011 * BF22 \quad (4.1)$$

dengan, $BF1 = (X_9 = 0)$;

$BF2 = (X_9 = 1)$;

$BF6 = (X_4 = 1 \text{ atau } X_4 = 2 \text{ atau } X_4 = 4 \text{ atau } X_4 = 5 \text{ atau } X_4 = 6) * BF2$;

$BF8 = \max(0, 193 - X_7) * BF6$;

$BF9 = (X_2 = 1)$;

$BF10 = (X_2 = 2)$;

$BF11 = (X_3 = 1) * BF9$;

$BF12 = (X_3 = 4 \text{ atau } X_3 = 5) * BF9$;

$BF13 = \max(0, X_7 - 292) * BF6$;

$BF15 = \max(0, X_7 - 250) * BF6$;

$BF17 = (X_4 = 2 \text{ atau } X_4 = 4) * BF12$;

$BF19 = (X_3 = 5) * BF10$;

$BF22 = \max(0, 240 - X_7) * BF19$;

Sebelum melakukan interpretasi pada model perlu dilakukan pengujian koefisien untuk mengetahui bahwa model MARS yang terpilih merupakan model yang sesuai dan menunjukkan hu-

bugan yang tepat antara variabel prediktor dengan variabel respon, berikut hipotesis yang digunakan untuk pengujian secara serentak.

$$H_0 : a_1 = a_8 = a_{11} = a_{13} = a_{15} = a_{17} = a_{22} = 0$$

$$H_1 : \text{minimal terdapat satu } a_j \neq 0, j = 1, 8, 11, 13, 15, 17, 22$$

Daerah Penolakan : Tolak H_0 apabila nilai $F_{\text{hitung}} > F_{\text{tabel}}$

Tabel 4.3. Statistik Uji Pengujian Serentak

F_{hitung}	F_{tabel} (0,05;7;61)	P-value	Keputusan
16,604	2,17	$4,80882 \times 10^{-12}$	Tolak H_0

Tabel 4.3 menunjukkan bahwa nilai F_{hitung} lebih besar dari F_{tabel} dan nilai P-value lebih kecil dari α (0,05) sehingga didapatkan keputusan Tolak H_0 yang artinya paling sedikit ada satu α_j yang signifikan dan model sesuai serta menunjukkan hubungan yang tepat antara variabel prediktor dan variabel respon. Selanjutnya dilakukan pengujian parsial (individu) pada Tabel 4.4 untuk mengetahui fungsi basis yang terbentuk didalam model, dengan hipotesis sebagai berikut.

$$H_0 : a_j = 0$$

$$H_1 : a_j \neq 0, j = 1, 8, 11, 13, 15, 17, \text{ dan } 22$$

Daerah Penolakan : Tolak H_0 apabila nilai $|T_{\text{hitung}}| > T_{\text{tabel}}$

Tabel 4.4. Statistik Uji Pengujian Parsial

Parameter	Estimasi	Standard Error	T_{hitung}	P-value
α_1	-1,000	0,106	-9,417	$0,168 \times 10^{-12}$
α_8	-0,004	0,001	-3,324	0,002
α_{11}	0,673	0,155	4,349	$5,28 \times 10^{-5}$
α_{13}	0,031	0,007	4,576	$2,38 \times 10^{-5}$
α_{15}	-0,029	0,005	-6,517	$1,56 \times 10^{-8}$

Tabel 4.4. (Lanjutan) Statistik Uji Pengujian Parsial

Parameter Estimasi	<i>Standard Error</i>	T_{hitung}	<i>P-value</i>	
α_{17}	0,253	0,069	3,658	$0,5312 \times 10^{-3}$
α_{22}	0,011	0,003	3,854	$0,281954 \times 10^{-3}$

Nilai Tabel T dengan $\alpha = 0,05$, dan Df = 61 didapatkan senilai 2,00, sehingga dari semua parameter didapatkan keputusan Tolak H_0 yang artinya terdapat pengaruh setiap variabel prediktor dengan variabel respon pada fungsi basis yang ada didalam model. Karena semua parameter telah signifikan pada model sehingga dapat dicari nilai peluang kejadian Infeksi Oportunistik (IO) menggunakan rumus (2.11) dengan model pada persamaan (4.1) sehingga didapatkan nilai perhitungan peluang kejadian IO sebagai berikut.

$$\hat{f}(x) = 0,993 - 1,000 * (1) - 0,004 * (1) + 0,673 * (1) + 0,031 * (1) - 0,029 * (1) + 0,253 * (1) + 0,011 * (1) = 0,928$$

sehingga,

$$\hat{\pi}(x) = \frac{e^{\hat{f}(x)}}{1 + e^{\hat{f}(x)}} = \frac{e^{0,928}}{1 + e^{0,928}} = \frac{2,52945}{3,52945} = 0,717 \text{ dan } 1 - \hat{\pi}(x) = 1 - 0,717 = 0,283$$

Jadi probabilitas tidak terjadinya kejadian Infeksi Oportunistik pada pasien adalah sebesar 0,717 dan terjadinya Infeksi Oportunistik adalah sebesar 0,283. Interpretasi model MARS pada persamaan 4.1 dan probabilitas kejadian IO untuk masing-masing fungsi basis dengan mengasumsikan konstan untuk beberapa fungsi basis lain secara bergantian adalah sebagai berikut.

1. $BF1 = (X_9 = 0)$;

Pasien HIV/AIDS yang mempunyai pengetahuan baik mempunyai peluang 0,498 tidak terjadinya kejadian Infeksi Oportunistik (IO) dan peluang terjadinya kasus Infeksi Oportunistik (IO) sebesar 0,502.

$$BF8 = \max(0, 193 - X_7) * BF6;$$

$$2. \quad BF6 = (X_4 = 1 \text{ atau } X_4 = 2 \text{ atau } X_4 = 4 \text{ atau } X_4 = 5 \\ \text{atau } X_4 = 6) * BF2;$$

$$BF2 = (X_9 = 1);$$

Pasien HIV/AIDS yang hanya mempunyai kadar CD4 kurang dari 193, dengan pendidikan yang tidak sekolah, atau pendidikan SD/ sederajat atau pendidikan SMA/ sederajat, atau Akademi DI hingga DIII atau Perguruan Tinggi S1 hingga S3, dengan pengetahuan yang kurang baik mempunyai peluang tidak terjadinya kasus IO sebesar 0,729 dan terjadinya kasus IO sebesar 0,271.

$$3. \quad BF11 = (X_3 = 1) * BF9;$$

$$BF9 = (X_2 = 1);$$

Pasien HIV/AIDS yang bekerja sebagai PNS dan berjenis kelamin laki-laki mempunyai peluang tidak terkena kasus Infeksi Oportunistik sebesar 0,841 dan terjadinya kasus Infeksi Oportunistik sebesar 0,159.

$$BF13 = \max(0, X_7 - 292) * BF6;$$

$$4. \quad BF6 = (X_4 = 1 \text{ atau } X_4 = 2 \text{ atau } X_4 = 4 \text{ atau } X_4 = 5 \\ \text{atau } X_4 = 6) * BF2;$$

$$BF2 = (X_9 = 1);$$

Pasien HIV/AIDS yang mempunyai kadar CD4 lebih dari 292, dengan pendidikan yang tidak sekolah, atau pendidikan SD/ sederajat atau pendidikan SMA/ sederajat, atau Akademi DI hingga DIII atau Perguruan Tinggi S1 hingga S3, dan pengetahuan yang kurang baik mempunyai peluang tidak terjadinya kasus IO sebesar 0,736 dan terjadinya kasus IO sebesar 0,264.

$$BF15 = \max(0, X_7 - 250) * BF6;$$

$$5. \quad BF6 = (X_4 = 1 \text{ atau } X_4 = 2 \text{ atau } X_4 = 4 \text{ atau } X_4 = 5 \\ \text{atau } X_4 = 6) * BF2;$$

$$BF2 = (X_9 = 1);$$

Pasien HIV/AIDS yang mempunyai kadar CD4 lebih dari 250, dengan pendidikan yang tidak sekolah, atau pendidikan SD/ sederajat atau pendidikan SMA/ sederajat, atau Akademi DI hingga DIII atau Perguruan Tinggi S1 hingga S3, dan pengetahuan yang kurang baik mempunyai peluang tidak terjadinya kasus IO sebesar 0,724 dan terjadinya kasus IO sebesar 0,276.

$$BF17 = (X_4 = 2 \text{ atau } X_4 = 4) * BF12;$$

$$6. \quad BF12 = (X_3 = 4 \text{ atau } X_3 = 5) * BF9;$$

$$BF9 = (X_2 = 1);$$

Pasien HIV/AIDS yang berpendidikan lulusan SD/ sederajat atau SMA/ sederajat, dengan pekerjaan Swasta atau Tidak bekerja dan berjenis kelamin laki-laki mempunyai peluang tidak terjadinya IO sebesar 0,777 dan terjadinya IO sebesar 0,223.

$$BF22 = \max(0, 240 - X_7) * BF19$$

$$7. \quad BF19 = (X_3 = 5) * BF10;$$

$$BF10 = (X_2 = 2);$$

Pasien HIV/AIDS yang mempunyai kadar CD4 kurang dari 240 dan tidak bekerja dengan jenis kelamin perempuan mempunyai peluang tidak terjadinya kasus IO sebesar 0,732 dan terjadinya kasus IO sebesar 0,268.

Selanjutnya ditunjukkan tingkat kepentingan variabel-variabel prediktor dalam mengklasifikasikan kejadian Infeksi Oportunistik menggunakan metode MARS, dan naiknya nilai GCV ditaksir karena berpindah-pindahnya variabel prediktor yang masuk dalam model.

Tabel 4.5 Tingkat Kepentingan Variabel Prediktor pada Model MARS

Variabel	Tingkat Kepentingan	Pengurangan Nilai GCV
X ₉	100.000	0.213
X ₇	68.661	0.180
X ₄	55.438	0.170
X ₂	37.375	0.159
X ₃	37.375	0.159

Tabel 4.5 (Lanjutan) Tingkat Kepentingan Variabel Prediktor

Variabel	Tingkat Kepentingan	Pengurangan Nilai GCV
X ₁	0.000	0.151
X ₅	0.000	0.151
X ₆	0.000	0.151
X ₈	0.000	0.151
X ₁₀	0.000	0.151
X ₁₁	0.000	0.151
X ₁₂	0.000	0.151

Tabel 4.5 menunjukkan bahwa terdapat lima variabel terpenting dalam pembentukan model dan masuk mempengaruhi klasifikasi kejadian IO. Variabel yang paling penting dan mempengaruhi pengurangan nilai GCV terbesar pada model adalah variabel pengetahuan pada pasien (X₉) dan apabila dimasukkan dalam model maka akan mengurangi nilai GCV terbesar sejumlah 0.213, variabel nilai CD4 akhir pada pasien HIV/AIDS yang terukur (X₇) apabila dimasukkan model maka akan mengurangi nilai GCV sebesar 0,180, variabel pendidikan (X₄) jika berada didalam model maka akan menurunkan nilai GCV sebesar 0,170, untuk variabel jenis kelamin (X₂) dan pekerjaan (X₃) apabila dimasukkan model maka masing-masing akan mengurangi nilai GCV sebesar 0,159. Untuk variabel-variabel lain dari X₁ hingga X₁₂ jika dimasukkan dalam model maka masing-masing akan mengurangi nilai GCV sebesar 0,151.

Tabel 4.6 Ketepatan Klasifikasi Data *Training* pada Model MARS untuk Kejadian Infeksi Oportunistik

Aktual Kejadian IO	Prediksi Kejadian Infeksi Oportunistik (IO)		Total Aktual
	Tdk ada IO (0)	Ada IO (1)	
0	53	1	54
1	4	11	15
Total Prediksi	57	12	69
TAR	92,80%	APER	7,20%
Sensitivity	98,10%	Specificity	73,30%

Dari sejumlah 69 data *training* terdapat 53 orang pasien yang terklasifikasi dengan benar untuk tidak adanya kejadian Infeksi Oportunistik (IO), dan 11 orang pasien terklasifikasi dengan benar untuk kejadian IO (kategori 1). Ketepatan klasifikasi yang dihasilkan adalah sebesar 92,80% dan kesalahan klasifikasi antara data aktual dan prediksi adalah sebesar 7,20%. Berdasarkan nilai *sensitivity* dapat diketahui bahwa terdapat 98,10% pasien HIV/-AIDS yang tepat diklasifikasikan tidak ada kejadian IO, sedangkan dari nilai *specificity* dapat diketahui bahwa terdapat 73,30% pasien HIV/AIDS yang tepat diklasifikasikan terjadi kasus IO. Nilai *sensitivity*, *specificity* dan *G-means* yang cukup besar dan berimbang menunjukkan bahwa metode MARS sesuai digunakan untuk data *training* dalam mengklasifikasikan kejadian IO. Selanjutnya dilakukan pengujian kestabilan dalam ketepatan klasifikasi kejadian IO dengan Hipotesis:

H_0 : Klasifikasi kejadian IO belum dianggap stabil dan konsisten secara statistik

H_1 : Klasifikasi kejadian IO dianggap sudah stabil dan konsisten secara statistik

Daerah Penolakan :

Tolak H_0 apabila nilai *Press'Q* lebih dari nilai Tabel $\chi_{\alpha,df}$

Tabel 4.7 Statistik Uji Kestabilan Data *Training* pada Model MARS

Ketepatan Klasifikasi	<i>Press'Q</i>	$\chi_{(0,05;1)}$
92,80%	50,44928	3,841

Tabel 4.7 menunjukkan hasil pengujian kestabilan klasifikasi dengan taraf signifikansi α sebesar 0,05 dan didapatkan nilai *Press'Q* sebesar 50,449 sehingga keputusan yang diambil adalah Tolak H_0 karena nilainya sudah lebih besar dari nilai tabel *chisquare* yang bernilai 3,841 yang artinya sudah stabilnya hasil klasifikasi kejadian IO pada data *training* dalam pemodelan MARS dan dapat dilanjutkan untuk mengklasifikasi data *testing*. Berdasarkan Tabel 4.8 dapat diketahui bahwa terdapat 11 dari 15 atau 73,33% pasien yang terklasifikasi benar tidak adanya kejadian Infeksi Oportunistik pada pasien dan ketepatan klasifikasi

pada data *testing* adalah sebesar 47,80% serta kesalahan klasifikasi yang lebih tinggi yakni sebesar 52,20%.

Tabel 4.8 Ketepatan Klasifikasi Data *Testing* pada Model MARS untuk Kejadian Infeksi Oportunistik

Aktual Kejadian IO	Prediksi Kejadian Infeksi Oportunistik (IO)		Total Aktual
	Tdk ada IO (0)	Ada IO (1)	
0	11	4	15
1	8	0	8
Total Prediksi	19	4	23
TAR	47,80%	APPER	52,20%
<i>Sensitivity</i>	73,30%	<i>Specificity</i>	0,00%
		<i>G-means</i>	0,00%

Pada nilai *sensitivity* menunjukkan bahwa terdapat 73,30% pasien HIV/AIDS yang tepat diklasifikasikan tidak ada IO, sedangkan nilai *specificity* dan *G-means* yang hanya bernilai 0% menunjukkan bahwa pada data *testing* metode MARS hanya sesuai untuk mengklasifikasikan kejadian tidak ada IO pada pasien HIV/AIDS dan data tidak seimbang.

Tabel 4.9 Statistik Uji Kestabilan Data *Testing* pada Model MARS

Ketepatan Klasifikasi	<i>Press'Q</i>	$\chi_{(0,05;1)}$
47,80%	0,0435	3,841

Dengan hipotesis yang sama pada data *training* sebelumnya dan taraf signifikansi α sebesar 0,05, Tabel 4.9 menunjukkan hasil bahwa masih belum stabilnya data *testing* untuk mengklasifikasi kejadian IO, karena nilai *Press'Q* yang dihasilkan kurang dari nilai tabel *chisquare*.

4.3 Klasifikasi Kejadian Infeksi Oportunistik pada Pasien HIV/AIDS Menggunakan Metode Bagging MARS

Metode Bagging MARS dilakukan dengan replikasi *bootstrap* pada data *training* dan data *testing* sebanyak 50, 100, 150, 200, dan 500, kali, sehingga didapatkan hasil prediksi pada variabel respon dan dapat dilakukan klasifikasi untuk mengetahui ketepatan klasifikasi kejadian Infeksi Oportunistik pada pasien

HIV/AIDS menggunakan metode Bagging MARS. Berikut hasil ketepatan klasifikasi pada data *training* dan data *testing* dengan berbagai kombinasi replikasi.

Tabel 4.10 Hasil Ketepatan Klasifikasi Data *Training* Metode Bagging MARS

Replikasi	Ketepatan Klasifikasi	Kesalahan Klasifikasi
50	92,80%	7,20%
100	91,30%	8,70%
150	92,80%	7,20%
200	92,80%	7,20%
500	92,80%	7,20%

Hasil ketepatan klasifikasi metode Bagging MARS yang ditunjukkan pada Tabel 4.10 menunjukkan bahwa dengan replikasi sejumlah 50 kali menghasilkan nilai ketepatan klasifikasi tertinggi sebesar 92,80% dan kesalahan klasifikasi terendah dengan nilai 7,20%. Jika nilai replikasi diulang beberapa kali lipat hasil ketepatan dan kesalahan klasifikasinya sudah stabil dan akan bernilai sama.

Tabel 4.11 Hasil Ketepatan Klasifikasi Data *Testing* Metode Bagging MARS

Replikasi	Ketepatan Klasifikasi	Kesalahan Klasifikasi
50	52,20%	47,80%
100	52,20%	47,80%
150	52,20%	47,80%
200	52,20%	47,80%
500	52,20%	47,80%

Sama halnya dengan hasil ketepatan klasifikasi untuk data *training* menggunakan metode Bagging MARS sebelumnya. Pada Tabel 4.11 hasil ketepatan klasifikasi data *testing* dapat diketahui bahwa replikasi sebanyak 50 kali juga mempunyai nilai ketepatan klasifikasi tertinggi yakni sebesar 52,20% dan kesalahan klasifikasi yang hanya berkisar 47,80% serta nilai ketepatan dan kesalahan klasifikasi yang konstan setelah nilai replikasi diperbesar. Estimasi parameter model Bagging MARS merupakan rata-rata nilai estimasi parameter B kali replikasi *bootstrap*, namun karena nilai knots untuk model Bagging MARS pada setiap kali replikasi *bootstrap* berubah-ubah maka untuk estimasi parameter-

nya tidak dapat dirata-rata, sehingga model untuk Bagging MARS tidak ada, dan metode Bagging MARS hanya digunakan untuk menurunkan kesalahan klasifikasi pada pemodelan MARS. Model yang digunakan untuk mengklasifikasikan kejadian Infeksi Oportunistik di Puskesmas Kabupaten Pasuruan adalah model MARS dengan $BF=24$, $MI=3$, dan $MO=1$.

4.4 Perbandingan Ketepatan Klasifikasi Kejadian IO Menggunakan Metode MARS dan Metode Bagging MARS

Untuk mendapatkan metode terbaik yang sesuai digunakan untuk mengklasifikasi kejadian IO antara metode MARS dan metode Bagging MARS dapat dilakukan dengan cara membandingkan besarnya nilai ketepatan klasifikasi atau akurasi dari kedua metode tersebut dengan hasil ditunjukkan pada Tabel 4.12 untuk data *training* dan tabel 4.13 untuk data *testing*.

Tabel 4.12 Perbandingan Hasil Ketepatan Klasifikasi Data *Training* Menggunakan Metode MARS dan Bagging MARS

MARS	Bagging MARS	
	Replikasi	Ketepatan Klasifikasi
92,80%*	50	92,80%
	100	91,30%
	150	92,80%
	200	92,80%
	500	92,80%

Perbandingan hasil ketepatan klasifikasi pada Tabel 4.12 menunjukkan bahwa pada data *training* didapatkan tingkat ketepatan klasifikasi yang sama antara metode MARS dan Bagging MARS yakni sebesar 98,20%. Nilai ketepatan yang tidak berubah karena dimungkinkan nilai tersebut merupakan hasil optimum data yang dapat diklasifikasikan sehingga untuk klasifikasi kejadian Infeksi Oportunistik pada data *training* sama-sama tepat menggunakan kedua metode tersebut. Hasil perbandingan keakuratan klasifikasi pada data *testing* menggunakan metode MARS dan Bagging MARS ditunjukkan pada Tabel 4.13 sebagai berikut.

Tabel 4.13 Perbandingan Hasil Ketepatan Klasifikasi Data *Testing* Menggunakan Metode MARS dan Bagging MARS

MARS	Bagging MARS	
	Replikasi	Ketepatan Klasifikasi
47,80%	50	52,20%
	100	52,20%
	150	52,20%
	200	52,20%
	500	52,20%

Tabel 4.13 menunjukkan perbandingan hasil ketepatan klasifikasi data *testing* menggunakan metode MARS dan Bagging MARS didapatkan hasil bahwa metode Bagging MARS lebih tepat digunakan untuk mengklasifikasikan kejadian Infeksi Oportunistik pada pasien HIV/AIDS, karena memiliki nilai ketepatan klasifikasi yang lebih besar dibandingkan metode MARS dengan nilai sebesar 52,20%.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil analisis dan pembahasan pada bab sebelumnya untuk hasil pemodelan dan ketepatan klasifikasi kejadian Infeksi Oportunistik (IO) pada pasien HIV/AIDS di Kabupaten Pasuruan, diperoleh kesimpulan sebagai berikut.

1. Karakteristik pasien HIV/AIDS menunjukkan bahwa 75% pasien sudah tidak ditemukan adanya Infeksi Oportunistik dan 25% sisanya masih terdapat adanya IO, dengan rata-rata umur pasien HIV/AIDS adalah berusia 33 tahun, bekerja swasta, dengan pendidikan SMA/ sederajat dan lama waktu kejadian penyakit HIV/AIDS berkisar selama 3 tahun serta rata-rata jumlah sel CD4 dalam darah pasien adalah sebesar 291,4 dan 79% diantaranya telah mendapat ARV dan pendampingan.
2. Dari model MARS yang telah terbentuk didapatkan probabilitas tidak terjadinya kasus Infeksi Oportunistik pada pasien sebesar 0,717 dan terjadinya IO adalah sebesar 0,283, sedangkan ketepatan klasifikasi kejadian IO pada data *training* yang dihasilkan adalah sebesar 92,80% dan untuk data *testing* adalah sebesar 47,80%.
3. Hasil ketepatan klasifikasi metode Bagging MARS untuk data *training* maupun data *testing* menunjukkan bahwa dengan replikasi sejumlah 50 kali menghasilkan nilai ketepatan klasifikasi tertinggi masing-masing sebesar 92,80% dan 52,20%, serta jika nilai replikasi diulang beberapa kali lipat hasil ketepatan dan kesalahan klasifikasinya sudah stabil dan akan bernilai sama.
4. Dari perbandingan kedua metode didapatkan kesimpulan bahwa pada klasifikasi kejadian IO dengan data *training* sama-sama tepat menggunakan kedua metode MARS maupun Bagging MARS, sedangkan untuk data *testing* lebih tepat jika menggunakan metode Bagging MARS.

5.2 Saran

Saran yang dapat diberikan untuk penelitian selanjutnya agar lebih memperhatikan adanya *missing* data dan menambah jumlah sampel data penelitian agar didapatkan model, peluang kejadian dan keakuratan klasifikasi yang lebih baik.

Sedangkan untuk pasien HIV/AIDS agar lebih patuh untuk melakukan terapi ARV dan untuk keluarga pasien, pihak pelayanan kesehatan Puskesmas, serta Dinas Kesehatan terkait agar lebih memperhatikan pasien HIV/AIDS sehingga dapat meningkatkan kualitas hidupnya dan memperhatikan faktor resiko terjadinya kasus Infeksi Oportunistik pada pasien HIV/AIDS.

DAFTAR PUSTAKA

- Arleina, O. D. (2014). *Bootstrap Aggregating Multivariate Adaptive Regression Splines (Bagging MARS) untuk Mengklasifikasikan Rumah Tangga Miskin di Kabupaten Jombang*. Surabaya: Tugas Akhir ITS.
- Breiman, L. (1994, September). Bagging Predictors. 421 (Technical Report), 17.
- Buhlmann, P., & Yu, B. (2002). Analyzing Bagging, The Annals of Statistics. Vol 30 No. 4, Hal 927-961.
- Cox, D., & Snell, E. (1989). *Analysis of Binary Data* (Second ed.). London: Chapman dan Hall.
- Departemen Kesehatan RI. (2010). *Riskesmas Tahun 2010*. Jakarta: Departemen Kesehatan RI.
- Dinas Kesehatan Provinsi Jawa Timur. (2016). *Profil Kesehatan Provinsi Jawa Timur 2015*. Surabaya: Dinkes Jatim.
- Eubank, R. (1999). *Spline Smoothing and Nonparametric Regression*. New York: Marcel Deker.
- Friedman, J. H. (1990). *Multivariate Adaptive Regression Spline*. California: Department of Statistics - Stanford University.
- Friedman, J., & Silverman, B. (1989). *Flexible Parsimony Smoothing and Additive Modelling: Technometrics* 31. Amerika.
- Hair JR., J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate Data Analysis* (Seventh ed.). New Jersey: Pearson Prentice Hall.
- Hasyim, M. (2010). *Pemodelan Angka Kejadian Penyakit Infeksi Tuberkulosis Paru (TB Paru) di Kabupaten Sorong Selatan (Provinsi Papua Barat) dengan Pendekatan Multivariate Adaptive Regression Spline (MARS)*. Surabaya: Tugas Akhir ITS.
- Iradat, D. (2015, November 30). *Jumlah Kasus HIV & AIDS di Indonesia Meningkat*. Dipetik Oktober 30, 2016, dari metrotvnews: <https://www.google.co.id/amp/www.metrotvnews>

- tvnews.com/ amp/akW05xBK-jumlah-kasus-hiv-aids-di-indonesia-meningkat.
- Musoke, R. (2007). *HIV/AIDS in Children. African Medical Research Foundation (AMREF)*. Kenya: Allan and Nesta Ferguson Trust.
- Nash, M., & Bradford, D. (2001). *Parametric and Non Parametric Logistic Regression for Prediction of Precense/Absence of an Amphibian*. Las Vegas, Nevada.
- Nasronuidin. (2012). *HIV/AIDS Pendekatan Biologi Molekuler, Klinis, dan Sosial*. Surabaya: Airlangga University press.
- Nidhomuddin. (2014). *Random Forest dan Multivariate Adaptive Regression Splines (MARS) Binary Response untuk Klasifikasi Penderita HIV/AIDS di Surabaya*. Surabaya: Tesis ITS.
- Otok, B. W., Guritno, S., Subanar, & Haryatmi, S. (2006). *Bootstrap dalam MARS untuk Klasifikasi Perbankan. Inferensi Jurnal Statistik* (Vol. 2, No.1). Surabaya: FMIPA ITS.
- Saputro, S. A., Melaniani, S., Wibowo, A., Wijanarko, B., & Sofro, M. A. (2013). *Pemodelan Mixture Survival Studi Kasus HIV/AIDS di Klinik VCT/CST RSUP Dr. Kariadi Semarang*. Surabaya: Airlangga University.
- Spiritia, Y. (1995). *Yayasan Spiritia*. Dipetik Maret 15, 2017, dari www.spiritia.or.id.
- Sudoyo, A. W., Setiyohadi, B., Alwi, I., K, M. S., & Setiati, S. (2007). *Buku Ajar Ilmu Penyakit Dalam* (Vol. 1). Jakarta: Departemen Ilmu Penyakit Dalam FKUI.
- Utami, I. P. (2016). *Pemodelan Resiko Kejadian Malnutrisi pada Pasien Anak Penderita Penyakit ISPA dengan Pendekatan MARS*. Surabaya: Unair.

LAMPIRAN

Lampiran 1. Surat Pernyataan Pengambilan Data



PEMERINTAH KABUPATEN PASURUAN
DINAS KESEHATAN

Jalan Raya Rad Km 15 Bangil Pasuruan Jawa Timur 67153
Telepon (0343) 748909 Fax 747919 email : dinkes@dinkes.pasurungkab.go.id

Pasuruan, 5 Mei 2017

Nomor : 443 / 1139 / 424.072/2017 Kepada Yth
Sifat : Penting 1. Kepala Puskesmas Prigen
Lampiran : - 2. Kepala Puskesmas Grati
Hal : Penelitian Data Penyakit

DI PASURUAN

Menindaklanjuti surat Kepala Departemen Statistika Institute Teknologi Sepuluh Nopember Surabaya, nomor 022674/ITS.VI.1.3/TU.00.09/2017 tanggal 11 April 2017, maka untuk keperluan penelitian data kesehatan di wilayah Kabupaten Pasuruan, dimohon bantuannya untuk menerima dan memberikan informasi terkait dengan data-data kesakitan khusus penderita HIV AIDS, TB da ISPA di wilayah kerja Saudara.

Adapun petugas yang mengambil data sebagai berikut :

1. Lyin Nahrifah
2. Widya UmROUTIN S
3. Kenny Cyntia L.G
4. Grininda Rahmadian P
5. Moch. Bagasaswara A.M

Demikian atas kerjasamanya disampaikan terima kasih.



Lampiran 2. Data Rekam Medis Pasien HIV/AIDS di Puskesmas
Kabupaten Pasuruan

No.	Y	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂
1	1	34	1	4	2	2	24	248	1	0	0	0	0
2	0	52	1	4	2	2	18	341	1	1	0	0	0
3	0	39	2	4	4	1	19	681	0	0	0	0	0
4	1	31	2	5	3	1	30	517	1	1	0	0	0
5	0	33	2	4	3	1	39	638	1	0	0	0	0
6	0	26	2	4	4	1	18	271	1	0	0	0	0
7	0	27	2	4	4	1	48	376	1	0	0	0	0
8	0	35	2	4	4	1	66	359	1	0	0	0	0
9	0	33	2	5	6	2	48	377	1	1	0	0	0
10	0	47	2	1	6	1	66	349	1	0	0	0	0
11	0	56	2	1	6	1	102	564	1	0	0	0	0
12	0	31	2	5	4	1	24	145	1	0	0	0	0
13	0	36	1	4	3	1	24	0	1	0	0	0	0
14	0	31	2	4	2	1	36	365	0	0	0	0	0
15	1	25	1	4	5	2	25	180	1	0	0	0	0
16	0	31	2	5	4	1	24	238	1	0	0	0	0
17	0	41	2	4	4	1	66	381	1	0	0	0	0
18	1	38	2	4	2	3	36	193	1	1	0	0	0
19	0	36	2	4	4	2	39	310	1	0	0	0	0
20	0	26	2	5	4	1	38	380	0	0	0	0	0
:	:	:	:	:	:	:	:	:	:	:	:	:	:
90	0	27	2	5	4	1	22	240	0	0	0	0	0
91	0	36	2	4	4	2	67	359	0	0	0	0	0
92	0	36	1	4	4	2	39	362	0	0	0	0	0

Lampiran 3. Model MARS Data *Training* dengan Kombinasi BF = 24, MI = 3, dan MO = 1 menggunakan AMOS 2.0

LEARNING SAMPLE STATISTICS

VARIABLE	MEAN	SD	N	SUM
Y	0.217	0.415	69.000	15.000
X1	33.594	9.323	69.000	2318.000
X2	1.536	0.502	69.000	106.000
X3	3.913	1.067	69.000	270.000
X4	3.899	1.214	69.000	269.000
X5	1.565	0.757	69.000	108.000
X6	38.899	29.857	69.000	2684.000
X7	283.406	159.552	69.000	19555.000
X8	0.783	0.415	69.000	54.000
X9	0.217	0.415	69.000	15.000
X10	0.000	0.000	69.000	0.000
X11	0.145	0.355	69.000	10.000
X12	0.043	0.205	69.000	3.000

Forward Stepwise Knot Placement

BasFn(s)	GCV	IndBsFns	EffPrms	Variable	Knot	Parent	BsF
0	0.175	0.0	1.0				
2 1	0.166	1.0	5.0	X9	10		
4 3	0.166	2.0	9.0	X5	00010	X9	2
6 5	0.175	3.0	13.0	X4	001000	X9	2
8 7	0.193	5.0	18.0	X7	193.000	X4	6
10 9	0.208	6.0	22.0	X2	10		
12 11	0.225	7.0	26.0	X3	100	X2	9
14 13	0.259	8.0	30.0	X7	292.000	X4	6
16 15	0.298	9.0	34.0	X7	250.000	X4	6
18 17	0.357	10.0	38.0	X4	010100	X3	12
20 19	0.456	11.0	42.0	X3	001	X2	10
22 21	0.566	13.0	47.0	X7	240.000	X3	19
24 23	0.819	14.0	51.0	X4	000100	X3	12

Final Model (After Backward Stepwise Elimination)

Basis Fun	Coefficient	Variable	Parent	Knot
0	0.993			
1	-1.000	X9		
8	-0.004	X7	X4	193.000
11	0.673	X3	X2	
13	0.031	X7	X4	292.000
15	-0.029	X7	X4	250.000
17	0.253	X4	X3	
22	0.011	X7	X3	240.00

Piecewise Linear GCV = 0.151, #efprms = 26.000

ANOVA Decomposition on 7 Basis Functions

fun	std. dev.	-gcv	#bsfns	#efprms	variable
1	0.412	0.315	1	3.571	X9
2	0.137	0.168	1	3.571	X2
3	0.317	0.183	3	10.714	X4
					X9
4	0.117	0.157	1	3.571	X2
					X4
5	0.120	0.160	1	3.571	X2
					X7

Piecewise Cubic Fit on 7 Basis Functions, GCV = 0.159
Relative Variable Importance

=====			
Variable		Importance	-gcv

9	X9	100.000	0.213
7	X7	68.661	0.180
4	X4	55.438	0.170
2	X2	37.375	0.159
3	X3	37.375	0.159
1	X1	0.000	0.151
5	X5	0.000	0.151
6	X6	0.000	0.151
8	X8	0.000	0.151
10	X10	0.000	0.151
11	X11	0.000	0.151
12	X12	0.000	0.151

ORDINARY LEAST SQUARES RESULTS

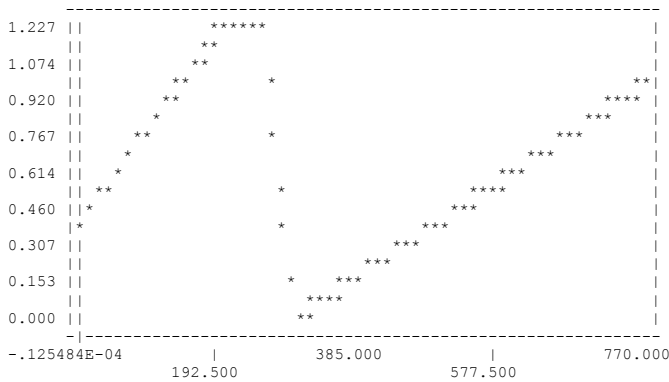
=====					
N: 69.000		R-SQUARED: 0.656			
MEAN DEP VAR: 0.217		ADJ R-SQUARED: 0.616			
		UNCENTERED R-SQUARED = R-0 SQUARED: 0.731			
PARAMETER		ESTIMATE	S.E.	T-RATIO	P-VALUE

Constant		0.993	0.099	10.066	.138778E-13
Basis Function 1		-1.000	0.106	-9.417	.168310E-12
Basis Function 8		-0.004	0.001	-3.324	0.002
Basis Function 11		0.673	0.155	4.349	.528341E-04
Basis Function 13		0.031	0.007	4.576	.238254E-04
Basis Function 15		-0.029	0.005	-6.517	.155549E-07
Basis Function 17		0.253	0.069	3.658	.531253E-03
Basis Function 22		0.011	0.003	3.854	.281954E-03

F-STATISTIC = 16.604		S.E. OF REGRESSION = 0.257			
P-VALUE = .480882E-11		RESIDUAL SUM OF SQUARES = 4.041			
[MDF,NDF] = [7, 61]		REGRESSION SUM OF SQUARES = 7.699			

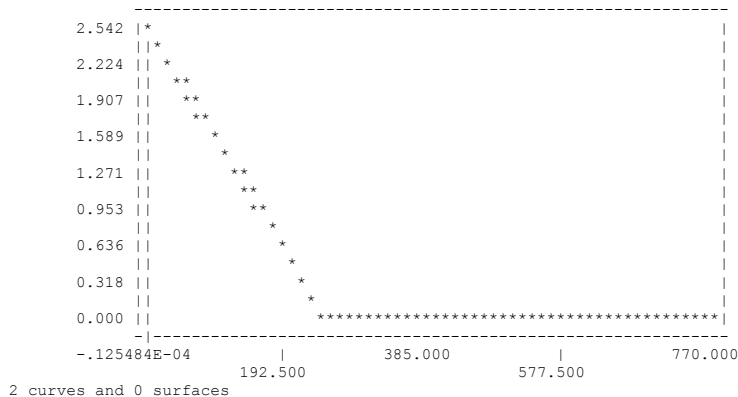
The Following Graphics Are Piecewise Linear
CATEGORICAL - ORDINAL INTERACTION:

X4 = 1 1 0 1 1 1
X9 = 0 1
CURVE 1: X7 , max = 1.2272



CATEGORICAL - ORDINAL INTERACTION:

X2 = 0 1
X3 = 0 0 1
CURVE 2: X7 , max = 2.5420



Basis Functions

```
=====
BF1 = ( X9 = 0 );
BF2 = ( X9 = 1 );
BF6 = ( X4 = 1 OR X4 = 2 OR X4 = 4 OR X4 = 5 OR X4 = 6 ) * BF2;
BF8 = max(0, 193.000 - X7 ) * BF6;
BF9 = ( X2 = 1 );
BF10 = ( X2 = 2 );
BF11 = ( X3 = 1 ) * BF9;
BF12 = ( X3 = 4 OR X3 = 5 ) * BF9;
BF13 = max(0, X7 - 292.000) * BF6;
BF15 = max(0, X7 - 250.000) * BF6;
BF17 = ( X4 = 2 OR X4 = 4 ) * BF12;
BF19 = ( X3 = 5 ) * BF10;
BF22 = max(0, 240.000 - X7 ) * BF19;
```

```
Y = 0.993 - 1.000 * BF1 - 0.004 * BF8 + 0.673 * BF11 + 0.031 * BF13
      - 0.029 * BF15 + 0.253 * BF17 + 0.011 * BF22;
```

model Y = BF1 BF8 BF11 BF13 BF15 BF17 BF22;

LEARNING SAMPLE CLASSIFICATION TABLE

=====			
Actual Class	Predicted Class		Actual Total
	0	1	
0	53.000	1.000	54.000
1	4.000	11.000	15.000

Pred. Tot.	57.000	12.000	69.000
Correct	0.981	0.733	
Success Ind.	0.199	0.516	
Tot. Correct	0.928		

```
Sensitivity:      0.981      Specificity:      0.733
False Reference:  0.070      False Response:    0.083
Reference = Class 0, Response = Class 1
```

Lampiran 4. Ketepatan Klasifikasi Data *Testing* dengan MARS

TEST SAMPLE CLASSIFICATION TABLE			
Actual Class	Predicted Class		Actual Total
	0	1	
0	11.000	4.000	15.000
1	8.000	0.000	8.000
Pred. Tot.	19.000	4.000	23.000
Correct	0.733	0.000	
Success Ind.	0.081	-0.348	
Tot. Correct	0.478		
Sensitivity: 0.733 Specificity: 0.000			
False Reference: 0.421 False Response: 1.000			
Reference = Class 0, Response = Class 1			

Lampiran 5. *Sintax* Program Bagging MARS pada *software* R untuk Replikasi Sejumlah 50 Kali

```
library(mda)
library(earth)
library(caret)
datatraining=read.table("e:/TERBAIKTRAINING.txt",
header=TRUE)
bagging.mars=bagEarth(Y~X1+X2+X3+X4+X5+X6+X7+X8,
data=datatraining, B=50)
bagging.mars.pred=predict(bagging.mars, data=datatraining)
write.csv(bagging.mars.pred,"e:\\bagmtra(50).csv")
datatesting=read.table("e:/TERBAIKTESTING.txt",
header=TRUE)
bagging.mars=bagEarth(Y~X1+X2+X3+X4+X5+X6+X7+X8,
data=datatesting, B=50)
bagging.mars.pred2=predict(bagging.mars, data=datatesting)
write.csv(bagging.mars.pred2,"e:\\bagmtes(50).csv")
```

Keterangan :

- Variabel Y yang digunakan merupakan Prediksi dari Model MARS
- Variabel X yang digunakan merupakan Fungsi Basis dari Model MARS

Lampiran 6. Ketepatan Klasifikasi Data *Training* dan *Testing* dengan Bagging MARS Sejumlah 50 kali Replikasi

Aktual * Prediksi50 Crosstabulation

		Prediksi50		Total	
		0	1		
Aktual	0	Count	53	1	54
		% of Total	76.8%	1.4%	78.3%
	1	Count	4	11	15
		% of Total	5.8%	15.9%	21.7%
Total	Count	57	12	69	
	% of Total	82.6%	17.4%	100.0%	

Aktual * Prediksi50 Crosstabulation

		Prediksi50		Total	
		0	1		
Aktual	0	Count	12	3	15
		% of Total	52.2%	13.0%	65.2%
	1	Count	8	0	8
		% of Total	34.8%	0.0%	34.8%
Total		Count	20	3	23
		% of Total	87.0%	13.0%	100.0%

Lampiran 7. *Trial and Error* Pembentukan Model MARS dengan Pembagian 80:20 pada Data

Nomor Model	Kombinasi			GCV	MSE	R ²	Keakuratan Klasifikasi
	BF	MI	MO				
1	24	1	0	0,18	0,166	0,306	0,767
2	24	1	1	0,181	0,134	0,465	0,836
3	24	1	2	0,184	0	0,233	0,233
4	24	1	3	0,18	0,166	0,306	0,767
5	24	2	0	0,184	0	0,233	0,233
6	24	2	1	0,167	0,15	0,374	0,822
7	24	2	2	0,184	0	0,233	0,233
8	24	2	3	0,184	0	0,233	0,233

Nomor Model	Kombinasi			GCV	MSE	R ²	Keakuratan Klasifikasi
	BF	MI	MO				
9	24	3	0	0,184	0	0,233	0,233
10	24	3	1	0,178	0,162	0,325	0,315
11	24	3	2	0,183	0,168	0,296	0,795
12	24	3	3	0,181	0,153	0,371	0,781
13	36	1	0	0,18	0,166	0,306	0,767
14	36	1	1	0,182	0,134	0,465	0,836
15	36	1	2	0,184	0	0,233	0,233
16	36	1	3	0,18	0,166	0,306	0,767
17	36	2	0	0,184	0	0,233	0,233
18	36	2	1	0,167	0,15	0,374	0,822
19	36	2	2	0,184	0	0,233	0,233
20	36	2	3	0,184	0	0,233	0,233
21	36	3	0	0,184	0	0,233	0,233
22	36	3	1	0,179	0,162	0,325	0,315
23	36	3	2	0,178	0,083	0,689	0,918
24	36	3	3	0,174	0,134	0,458	0,822
25	48	1	0	0,181	0,166	0,306	0,767
26	48	1	1	0,184	0	0,233	0,233
27	48	1	2	0,184	0	0,233	0,233
28	48	1	3	0,181	0,166	0,306	0,767
29	48	2	0	0,184	0	0,233	0,233
30	48	2	1	0,167	0,15	0,374	0,822
31	48	2	2	0,184	0	0,233	0,233
32	48	2	3	0,184	0	0,233	0,233
33	48	3	0	0,184	0	0,233	0,233
34	48	3	1	0,179	0,162	0,325	0,315
35	48	3	2	0,178	0,083	0,689	0,918
36	48	3	3	0,174	0,134	0,458	0,822

Lampiran 8. *Trial and Error* Pembentukan Model MARS
dengan Pembagian 90:10 pada Data

Nomor Model	Kombinasi			GCV	MSE	R ²	Keakuratan Klasifikasi
	BF	MI	MO				
1	24	1	0	0,189	0	0,244	0,244
2	24	1	1	0,189	0	0,244	0,244
3	24	1	2	0,189	0	0,244	0,244
4	24	1	3	0,183	0,17	0,321	0,768
5	24	2	0	0,181	0,166	0,335	0,317
6	24	2	1	0,172	0,159	0,364	0,805
7	24	2	2	0,182	0,166	0,335	0,317
8	24	2	3	0,182	0,166	0,335	0,317
9	24	3	0	0,18	0,166	0,335	0,317
10	24	3	1	0,171	0,124	0,523	0,841
11	24	3	2	0,173	0,098	0,636	0,866
12	24	3	3	0,18	0,141	0,452	0,585
13	36	1	0	0,183	0,17	0,321	0,768
14	36	1	1	0,189	0	0,244	0,244
15	36	1	2	0,183	0,17	0,321	0,768
16	36	1	3	0,183	0,17	0,321	0,768
17	36	2	0	0,181	0,166	0,335	0,317
18	36	2	1	0,172	0,159	0,364	0,805
19	36	2	2	0,181	0,166	0,335	0,317
20	36	2	3	0,183	0,17	0,321	0,768
21	36	3	0	0,181	0,166	0,335	0,317
22	36	3	1	0,171	0,124	0,523	0,841
23	36	3	2	0,173	0,139	0,46	0,78
24	36	3	3	0,182	0,166	0,335	0,317

Nomor Model	Kombinasi			GCV	MSE	R ²	Keakuratan Klasifikasi
	BF	MI	MO				
25	48	1	0	0,183	0,17	0,321	0,768
26	48	1	1	0,189	0	0,244	0,244
27	48	1	2	0,183	0,17	0,321	0,768
28	48	1	3	0,189	0	0,244	0,244
29	48	2	0	0,181	0,166	0,335	0,317
30	48	2	1	0,172	0,159	0,364	0,805
31	48	2	2	0,181	0,166	0,335	0,317
32	48	2	3	0,183	0,166	0,335	0,317
33	48	3	0	0,181	0,166	0,335	0,317
34	48	3	1	0,171	0,124	0,523	0,841
35	48	3	2	0,173	0,139	0,46	0,78
36	48	3	3	0,182	0,166	0,335	0,317

BIODATA PENULIS



Penulis bernama lengkap Lyyin Nahriyah lahir di kota Lamongan pada tanggal 24 Agustus 1994. Anak bungsu dari dua bersaudara pasangan H. Toha Faqih dan Hj. Nurul Maghfiroh. Pendidikan formal yang ditempuh penulis antara lain TK Tashlihus Shibyan, SDN Duduklor, SMP Negeri 42 Surabaya, SMA Negeri 18 Surabaya, dan Diploma III Statistika ITS. Pada tahun 2015, penulis melanjutkan ke program Sarjana di Jurusan Statistika ITS dan lulus pada tahun 2017, dengan menyelesaikan Tugas Akhir yang berjudul “***Bootstrap Aggregating Multivariate Adaptive Regression Splines (Bagging MARS)* untuk Klasifikasi Pasien HIV/AIDS di Kabupaten Pasuruan**”. Selama masa Diploma hingga Sarjana, penulis pernah menjadi Staf Divisi PSt periode tahun 2013/2014, semifinal NSC tahun 2014 di Universitas Brawijaya, dan beberapa kegiatan surveyor. Bagi pembaca yang memiliki saran, kritik atau ingin berdiskusi lebih lanjut dengan penulis terkait dengan metode pada Tugas Akhir ini maupun keilmuan statistik dan hal-hal yang membuka wawasan lainnya dapat menghubungi email : lyyin.nahriyah@gmail.com.